心理科学进展 2021, Vol. 29, No. 11, 1948–1969
Advances in Psychological Science

https://doi.org/10.3724/SP.J.1042.2021.01948

·研究方法(Research Method)·

密集追踪数据分析：模型及其应用*

郑舒方 张沥今 乔欣宇 潘俊豪

(中山大学心理学系, 广州 510006)

摘 要 在心理学、教育学和临床医学等领域, 越来越多的研究者开始关注个体内部的行为、心理、临床效果等随时间而产生的动态变化, 重视针对个体的差异化建模。密集追踪是一种在短时间内对个体进行多个时间节点密集追踪测量的方法, 更适合用于研究个体内部心理过程等的动态变化及其作用机制。近年来, 密集追踪成为心理学研究的一大热点, 但许多密集追踪的研究分析仍停留在较为传统的方法。方法学领域已涌现出较多用于密集追踪数据分析的模型方法, 较为主流的模型包括以动态结构方程模型(Dynamic Structural Equation Model, DSEM)为代表的自上而下的建模方法, 以及以组迭代多模型估计(Group Iterative Multiple Model Estimation, GIMME)为代表的自下而上的建模方法。二者均可以方便地对密集追踪数据中的自回归及交叉滞后效应进行建模。

关键词 密集追踪, 时间序列, 动态结构方程模型, 组迭代多模型估计
分类号 B841

1 引言

在干预治疗或自然发生的事件中, 个体的想法、情绪、行为以及生理功能等往往不是一成不变的, 而是会随时间的推移发生动态变化(Vallacher et al., 2002)。传统的追踪研究或横断研究方法能够在一定程度上对人群中存在的现象进行描述或解释, 但难以揭示短时间内个体内部心理状态的动态变化, 也不能详细解释事件对个体的情绪、行为等产生的持续性影响(Setodji et al., 2019), 可能导致研究者对干预效果及其潜在机制产生错误解读(张银普 等, 2016; Setodji et al., 2019)。因此, 为了更深入地从个体内的角度研究变量间的关系, 研究者往往需要采用密集追踪方法, 更密集地测量个体状态发展和变化的细节(Trull & Ebner-Priemer, 2013)。

密集追踪数据有测量间隔短、频次高的特点, 常常被用来对个体的行为以及状态随时间变化的动态过程进行更精细的测量, 测量内容包括情绪、人格、认知和行为模式等方面(Hofmans et al., 2019; Lazarevic et al., 2020; Trull & Ebner-Priemer, 2014)。根据 Bringmann 等人(2017)的定义, 密集追踪数据是指研究者在较短的时间间隔内, 收集到的个体在不同时间点的变量信息。通过对变量进行重复测量, 研究者可以直接获取有关被试内的变量信息, 实时地记录被试随时间变化的动态过程(Bolger & Laurenceau, 2013), 进一步解释个体内部心理状态和行为的动态变化等(张银普 等, 2016)。

近年来, 随着手机等移动通讯设备的发展和使用, 研究者可以更加方便、频繁地收集有关被试行为、所处环境、心理和生理状态等方面的数据; 数据收集过程的便捷化以及收集场景的多样化等, 也都有利于研究者探究关于不同场景和不同时间点的问题。如今密集追踪数据已经被广泛应用到心理学的多个领域(Chen & Zhang, 2020; Windt et al., 2018), 如, 成瘾行为(Lanza et al., 2014; Shiffman, 2009; Weinstock et al., 2017)、压力症状(Chun, 2016; van de Schoot et al., 2015)、心理资本(Malmberg & Martin, 2019; Sturgeon et al.,

收稿日期: 2020-08-25
* 国家自然科学基金项目(31871128); 教育部人文社会科学研究规划基金项目(18YJA190013)。
通信作者: 潘俊豪, E-mail: panjunh@mail.sysu.edu.cn

2014)、饮食障碍问题(Munsch et al., 2009)、亲密关系(Howland & Rafaeli, 2010)、人格障碍(Trull et al., 2008; Wright et al., 2015, 2016)和语言使用(Lazarevic et al., 2020)等。

Trull 和 Ebner-Priemer (2013)指出, 密集追踪方法解决了很多回顾性调查和实验室方法的局限, 如潜在的记忆偏差、对个体日常生活场景的还原度不足等(Bolger & Laurenceau, 2013; Sened et al., 2018)。考虑到密集追踪方法存在的诸多优势, 国内一些研究者开始使用密集追踪数据来验证实验假设, 如探究干预方法对个体心理和行为变化的影响(陈明瑞, 周萍, 2017; 郑文倩 等, 2019)、环境因素对个体身心状态的影响(安媛媛, 徐慰, 2017)、个体在日常生活中的压力应对(邵华 等, 2019)以及个体在休闲活动中主观幸福感的变化等(任杰 等, 2010; 张峰, 耿晓伟, 2015)。这些研究多采用较为传统的多层线性模型(Multilevel Model, MLM)对密集追踪数据进行分析(安媛媛, 徐慰, 2017; 邵华 等, 2019; 张银普 等, 2016)。然而, 密集追踪方法的优势更多地在于探究变量在变化过程中的动态效应(即变量随时间的变化可能受到变量本身或其他外部因素的影响; Hamaker & Wichers, 2017; Molenaar, 2004; Molenaar & Campbell, 2009), 因此对密集追踪数据的分析方法提出了更高的要求。

为了帮助研究者们增进对于密集追踪数据方法的了解, 下文将首先对常见的密集追踪研究方法及其优势进行整体梳理, 然后介绍密集追踪数据中较为常用的分析方法(MLM), 以及目前较新的分析方法, 主要包括动态结构方程模型(Dynamic Structural Equation Model, DSEM)和组迭代多模型估计(Group Iterative Multiple Model Estimation, GIMME)。为了更直观地比较 MLM、DSEM 和 GIMME 在实际建模过程中的特点, 本文分别利用这三种模型方法进行实证的演示和比较。最后, 为了帮助研究者针对自身研究的实际情况灵活选择适合的模型方法, 本文还将对各种主流模型方法的优缺点和适用范围进行总结, 以供参考。

2　密集追踪研究方法及优势

2.1　密集追踪研究方法

密集追踪方法通常又称动态评估法(Ambulatory Assessment, AA; Bolger et al., 2003; Sened et al., 2018), 其另外几种常见的术语有：(1)经验取样法(Experience Sampling Method, ESM; Csikszentmihalyi & Larson, 2014), 广泛应用于人格和社会心理学领域, 主要使用主观问卷进行测量; (2)日记法(Daily Diary; Bolger & Laurenceau, 2013)和日常过程研究法(Daily Process Approach; Tennen et al., 2000), 通过被试的每日记录, 追踪有关变量随时间变化的过程; (3)生态瞬时评估法(Ecological Momentary Assessment, EMA; Stone & Shiffman, 1994), 常用于临床和药物研究。典型的 EMA 研究收集被试在日常生活中每日特定时刻的状态和行为数据, 在这些特定的应答时刻中, 被试或是被动接受来自外界的提示, 或是自主发起某些行为(如, 被试需要在感受到对成瘾药物的强烈欲望后进行作答; Lazarevic et al., 2020); (4)动态监测法(Ambulatory Monitoring; e.g., O'Brien et al., 2000), 被试需要佩戴特定的便携式仪器, 对特定的检测指标(如血压)进行密集测量。

一般来说, 密集追踪研究方法可以被广义地分为三类(Fisher & March, 2012; Trull & Ebner-Priemer, 2013), 包括自我报告式的自我评估; 观察式的自我评估; 生理、心理或行为的自我评估。在自我报告式的评估中, 根据不同的实验目的, 被试可以在固定/随机间隔的时间点进行自我评估, 或者在特定事件发生后进行作答(Reis et al., 2014)。不同的自我报告方式各有优势和不足：(1)固定时间间隔的自我报告方式要求被试在约定好的若干时间点回答相应的问题(Fisher & March, 2012), 其数据收集的过程不会突然打扰到被试, 数据收集方式本身简单易操作。然而, 固定时间间隔的自我报告也可能因为只考虑固定时间点发生的事件, 而忽略了其他时间点的事件可能产生的影响; 同时, 当测量时间间隔过大, 也可能使得被试出现记忆衰退的现象(Fisher & March, 2012; Hofmans et al., 2019)。(2)在随机时间间隔的自我报告中, 被试需要不定期地配合研究者完成数据的收集(Fisher & March, 2012), 因此可以随机抽到一些具有代表性的经历; 但是对于被试而言, 每次数据收集的时间是未知的, 这种不确定性可能对被试的日常生活带来较大的影响和负担(Fisher & March, 2012; Hofmans et al., 2019)。(3)关注特定事件的数据收集方式, 是指被试在发生

特定事件后主动配合研究者完成数据的收集(Fisher & March, 2012), 其优势在于可以收集到与研究者感兴趣的事件相关的数据。然而, 这种方法收集到的数据取决于被试对事件的反应, 因此研究者需要事先对被试进行训练, 使其能意识到特定事件的发生并主动及时向研究者报告(Fisher & March, 2012; Hofmans et al., 2019)。

另外, 与自我报告的评估不同, 观察性自我评估和生物、心理或行为的自我评估主要收集客观数据和指标(如, 被试所处环境中的声音、被试运动的速度, 以及心率等生理方面的指标数据), 而不是主观报告的数据(Trull & Ebner-Priemer, 2013), 因此很少使被试产生心理负担, 同时其数据收集也不易受到被试的主观影响。具体关于不同密集追踪方法的使用可以参考 Fisher 和 March (2012)、van Roekel 等人(2019)的文章, 以及本文关于密集追踪研究中的实际研究示例。

2.2 密集追踪研究方法的优势

相比于传统横断研究的方法, 密集追踪研究方法的优势主要体现在: (1)与横断研究中回顾式自我报告的方式相比, 密集追踪方法对变量的测量更及时、更接近被试的真实经历, 可以更好地减少测量过程中出现的记忆偏差(Stone et al., 1999)。因此, 所产生的密集追踪数据对于事件的描述往往更加准确, 更能代表被试的真实反应(Sened et al., 2018)。(2)与横断研究中的实验室方法相比, 密集追踪方法提供了多个时间点和场景的测量, 使得研究者可以探究个体内部心理构念随时间的动态发展及相互作用(张银普 等, 2016; Trull & Ebner-Priemer, 2014)。(3)密集追踪方法的测量随着被试的日常生活情境进行, 不仅局限于实验室环境, 极大降低了社会期望以及被试自我暴露意愿对实验数据的影响, 因此有更好的生态效度和外部效度(Sened et al., 2018)。(4)密集追踪方法对多变量重复测量, 允许研究者分别探究个体间和个体内的变异(Curran & Bauer, 2011; Sened et al., 2018)。许多心理学研究的理论和假设主要针对个体内部的过程(Intraindividual Processes), 但其数据分析的本质往往是对个体间关系(Interindividual Processes)的探讨(如, Lau 等人(2016)针对个体间自变量水平变化对因变量的影响, 得出"孕妇对压力的积极感知会促进其积极应对的策略, 进而提高身心健康"这一个体内的结论), 因此这样得到的关于个体内关系的结论可能是存在问题的(Curran & Bauer, 2011)。

相比于传统的追踪方法, 密集追踪也具有其独特的优势。传统追踪方法收集的数据重复观察的数量较少(通常少于 10 次; Jahng, 2008), 并且测量时间间隔较大, 因此主要用来探究在相对宏观的时间尺度上, 因变量总体均值随时间的变化(McNeish & Hamaker, 2019; Nesselroade, 1991; Setodji et al., 2019), 但不适合用于研究个体内变量的动态变化过程(Jahng, 2008)。而密集追踪方法对变量的测量间隔和持续时间更短、频次更多(Hamaker & Wichers, 2017; Walls & Schafer, 2006), 可以更深入地反映变量随时间变化的复杂模式(Bolger & Laurenceau, 2013), 允许研究者对被试的行为或心理社会过程随时间产生的细微变化, 以及相关的共变机制等进行研究(Collins, 2006)。相比于传统的追踪方法, 密集追踪方法在被试内水平的数据点较多, 有足够的检验力检验被试内水平的研究假设, 改善了传统追踪方法中被试内水平的数据点较少而统计检验力不足的问题(Fraley & Hudson, 2014; Hofmans et al., 2019)。

总的来说, 密集追踪研究方法能更好地保证数据的有效性和真实性(Sened et al., 2018), 允许研究者从个体的角度探究变量的变化关系, 深入分析个体内部的心理变化过程(Baltes & Nesselroade, 1979; Molenaar, 2004)。利用密集追踪方法, 研究者可以建立更有针对性的、针对于个体的模型, 从而更加精准地识别个体特征并对其进行更有效的治疗(Wright & Woods, 2020); 并通过探究变量在被试间水平的差异, 检验单个个体的效应是否能够推广或者泛化到更多的样本中(Hofmans et al., 2019)。

3 密集追踪数据的传统分析方法

密集追踪数据具有典型的分层结构, 即水平 1(时间点水平)的基本单元嵌套于水平 2(个体水平)的结构, 水平 2 的结构可以继续嵌套于更高的水平(如, 组别水平), 因此传统上利用 MLM 进行分析(张银普 等, 2016; Schwartz & Stone, 2007; Walls & Schafer, 2006), 将变量的变异来源分为个体内和个体间等多个水平(Walls & Schafer, 2006)。

3.1 方法原理

以两层的 MLM 为例, 一般来说, 密集追踪数

据中的水平 1 关注个体内变量在不同时间点之间的关系，而水平 2 关注个体水平上的变量间的关系。对于水平 1(时间点水平)，以存在一个水平 1 预测变量的情况为例，个体 i 在测量时间点 t 的模型公式如下(Heck & Thomas, 2015)：

$$Y_{it} = \beta_{0i} + \beta_{1i}X_{it} + \varepsilon_{it} \tag{1}$$

其中，X_{it} 表示个体 i 在第 t 个测量时间点的测量，β_{0i} 和 β_{1i} 分别表示截距和斜率，ε_{it} 表示误差项($\varepsilon_{it} \sim N(0,\sigma^2)$)。另外，在密集追踪数据的模型中，时间 T_{it} 可以作为预测变量加入到水平 1 的模型中。

对于水平 2(个体水平)，同样以存在一个水平 2 预测变量的情况为例，其模型表达式如下：

$$\beta_{0i} = \gamma_{00} + \gamma_{10}W_i + u_{0i} \tag{2}$$

$$\beta_{1i} = \gamma_{01} + \gamma_{11}W_i + u_{1i} \tag{3}$$

其中，β_{0i} 和 β_{1i} 分别为水平 1 中的随机截距和随机斜率，W_i 为水平 2 的预测变量，γ_{10}、γ_{11} 分别为水平 2 的预测变量 W_i 对 β_{0i}、β_{1i} 的斜率，γ_{00}、γ_{01} 表示对应的截距。各变量在水平 1 的斜率和截距可能在不同个体间存在变异(随机效应)，且这一变异可能可以被水平 2 的预测变量解释。另外，水平 1 和水平 2 的模型均可以处理含有多个预测变量的情况。根据实际研究中密集追踪数据的具体情况，可以对其进行更多层的建模，如测量时间点嵌套于个体，个体进一步嵌套于所在的群体等。

在使用 MLM 分析传统的追踪数据时，通常假设不同个体在水平 1 的残差方差 σ^2 均相等，但在密集追踪数据中，这种假设很可能会带来一些问题(Jahng & Wood, 2017)，如参数标准误的有偏估计等(Hamaker & Grasman, 2015)。密集追踪数据中多次重复测量的时间间隔较短，个体在当前时间点的测量很可能会受到先前测量的影响，因此通常需要假设水平 1 的残差带有自回归的效应(Hamaker & Grasman, 2015)。以考虑一阶自回归效应(Lag-1 Autoregressive Effects)为例，此时式(1)中的残差 ε_{it} 可以表示为(Hamaker & Grasman, 2015)：

$$\varepsilon_{it} = \Phi_i\varepsilon_{i,t-1} + e_{it} \tag{4}$$

其中，残差项 e_{it} 可以视为模型中没有考虑到的因素的影响，服从正态分布，即 $e_{it} \sim N(0, \sigma_e^2)$。$\Phi_i$ 表示个体 i 的自回归效应，即该个体在当前测量时间点的测量偏差 $_{it}$ 在多大程度上受到上一测量时间点的影响。自回归系数 Φ_i 在水平 2 的表达式为(Hamaker & Grasman, 2015)：

$$\Phi_i = \Phi + u_{2i} \tag{5}$$

其中，Φ 表示所有个体平均的自回归效应，u_{2i} 表示个体 i 的自回归效应与平均自回归效应之间的偏移，服从正态分布，$u_{2i} \sim N(0, \sigma_{u2}^2)$。

3.2　传统多层线性模型的不足

传统的 MLM 可以解释不同个体间变量随时间变化的关系曲线存在的差异(Walls & Schafer, 2006)，但是这种模型不够灵活。在分析动态变化效应、探究变量间关系的强度和方向随时间变化的趋势时，传统的 MLM 仍存在一定的局限(Li et al., 2006)。首先，当研究问题关注的是个体内心理过程的动态变化，如变异性和不稳定性时，传统的 MLM 方法在分析个体内发展模式中的个体间差异时可能会产生不正确的参数估计(Jahng, 2008)，且不能解决变量随时间变化的轨迹在个体间存在较大差异的问题(Shiyko et al., 2012)。其次，MLM 通常假设随机效应服从一定的分布类型，如正态分布；但通常情况下，这种假设很难得到满足，因此基于这些假设得到的模型参数也可能是不准确的(Piccirillo & Rodebaugh, 2019)。针对这一问题，Ke 等人(2019)指出，将贝叶斯多层结构方程模型(Bayesian Multilevel Structural Equation Modeling)与更稳健的、使用非正态分布(如 Student's t 分布，Student's t Distribution；混合正态分布，Mixture of Normal Distributions; Tong & Zhang, 2012; Zhang, 2013; Zhang et al., 2013)的贝叶斯估计方法相结合，可能可以解决随机效应不符合正态分布时参数估计不准确的问题。

另外，在 MLM 中，研究者往往需要通过中心化来分解组间水平和组内水平的效应(Curran & Bauer, 2011)。传统上对于预测变量的处理方法包括个体均值中心化、总均值中心化等，但这些方法的适用范围较为有限，在一些情况下容易产生较大的偏差(Asparouhov & Muthén, 2019a; Curran & Bauer, 2011)。如，在存在情境效应[1] (Contextual Effects; Lüdtke et al., 2008)或自回归效应(Nickell, 1981)，以及预测变量为(两个及以上类别的)分类变量且存在非线性连接函数(Link Function)的情

[1] 即，在控制了水平 1 的自变量 X_{ij} 的情况下，水平 2 的平均水平 $\bar{X}_j$ 仍对因变量存在额外的效应。

况下，传统方法均会产生较大的参数估计偏差(Asparouhov & Muthén, 2019a)。当缺失数据是随机缺失(Missing at Random, MAR)而不是完全随机缺失(Missing Completely at Random, MCAR)时，MLM 中的样本均值无法体现预测变量真实的均值，传统的中心化处理同样可能带来参数估计上的偏差(Asparouhov & Muthén, 2019a)。

4 密集追踪数据分析方法的新进展

为了克服传统方法在密集追踪数据中的局限性，帮助研究者更好地挖掘个体内部动态变化的过程和作用机制，下面将主要介绍两种较新的密集追踪数据的分析方法，包括 DSEM (Asparouhov et al., 2018)和结合了个体和群体信息的 GIMME (Gates & Molenaar, 2012)。根据 Hamaker 等人(2018)的界定，多个体的密集追踪数据分析方法可以分为自上而下的方法(Top-Down Approach)和自下而上的方法(Bottom-Up Approach)两大类。自上而下的方法通过较高水平的模型向下对较低水平的动态效应进行解释(如 MLM，DSEM 等)，这里的高低次序均为相对的而不是绝对的高低(如在三水平的 MLM 中，水平 2 相对于水平 1 是较高水平，但相对水平 3 是较低的水平)。而自下而上的方法首先对每个个体的数据进行分析，再进一步向上搜寻不同个体间相似的模式(如 GIMME)。

4.1 自上而下的方法：动态结构方程模型

DSEM 承接了传统 MLM 自上而下的思想，同时，由于它采用了贝叶斯的框架，因此相比传统 MLM 更为灵活(Asparouhov et al., 2018)。DSEM 结合了多种模型方法的优势，包括 MLM、时间序列模型(Time-Series Modeling)、结构方程模型(Structural Equation Modeling, SEM)以及时变效应模型(Time-Varying Effects Modeling, TVEM)，被认为是用于分析密集追踪数据最有前景的方法之一(Zhou et al., 2020)。它强调变量在时间尺度上的动态变化过程，这里的“动态”可以理解为变量在当前时刻的测量会受到变量自身在先前时间点的影响(自回归效应)，以及受到其他变量在先前(交叉滞后效应，Cross-lagged Effects)以及当前时间点(同时效应，Contemporaneous Effects)的测量的影响。另外，在 DSEM 中，各变量的变异均可通过个体特异和时间特异的变量进行解释。

4.1.1 方法原理

DSEM 可以认为是动态因子模型(Dynamic Factor Model, DFM; Ferrer & Nesselroade, 2003; Molenaar, 1985; Molenaar et al., 1999; Zhang & Nesselroade, 2007)在多个个体情况下的扩展(Asparouhov et al., 2018)。它结合多层模型的思路，将第 i 个个体在第 t 个时间节点的观测变量 Y_{it}，分解为个体 i 的特异贡献 $Y_{2,i}$、时间 t 的特异贡献 $Y_{3,t}$，以及除去个体和时间的特异贡献后，个体 i 在时间节点 t 的波动变化 $Y_{1,it}$ 这三个成分，即 $Y_{it}=Y_{1,it}+Y_{2,i}+Y_{3,t}$ (Asparouhov et al., 2018)。例如，对被试群体在工作日中感受到的工作压力和积极情绪水平进行测量，其中个体 i 在时间节点 t 的工作压力水平 Y_{it}，将同时取决于个体 i 在工作压力上的一般基线水平 $Y_{2,i}$、时间节点 t 对个体工作压力的一般效应 $Y_{3,t}$ (如，一般来说，每到周一时个体的工作压力都会增高)，以及除去个体和时间的一般性效应后，个体 i 在时间节点 t 的工作压力变化 $Y_{1,it}$ (如公式(6)~(7)所示，这些变化的来源可能包括：该个体在该时间节点的积极情绪变化，先前时间点的积极情绪和工作压力水平对当前时间点的影响等)。

最完整的 DSEM 模型的框架包括 $Y_{2,i}$、$Y_{3,t}$ 以及 $Y_{1,it}$ 这三个成分，但不是所有的密集追踪研究均存在时间特异的效应，因此不是所有的密集追踪研究均适合采用最完整的 DSEM。如，在密集追踪的研究中，如果 $t=1$ 仅表示个体的第一个观测时间点的测量，而不是某个特定的时间节点(比如 $t=1$ 都表示每周的第一天)时，$t=1$ 对于每个个体来说没有一致的含义，因此数据本身很可能不存在时间特异的效应(Asparouhov et al., 2018)。Asparouhov 等人(2018)指出，研究者可以根据实际研究的需要而舍弃 $Y_{3,t}$ 这一成分，只考虑个体水平的随机效应，即采用两层的 DSEM (Two-Level DSEM)；甚至进一步舍弃 $Y_{2,i}$，只考虑单层的 DSEM 模型(Single-Level DSEM)。两层和单层的 DSEM 均为总框架下的特殊的情况，因此，本文将对总的 DSEM 框架下的方法原理进行介绍。

首先，与 DFM 类似，DSEM 在水平 1 考虑了变量在时间节点 t 的测量可能受到当前及先前 L 次测量的影响，包括测量模型和结构方程模型(Asparouhov et al., 2018)：

$$Y_{1,it}=\nu_1+\sum_{l=0}^{L}\Lambda_{1,lit}\eta_{1,i,t-l}+\sum_{l=0}^{L}R_{lit}Y_{1,i,t-l}+\sum_{l=0}^{L}K_{1,lit}X_{1,i,t-l}+\varepsilon_{1,it} \quad (6)$$

$$\eta_{1,it}=\alpha_{1,it}+\sum_{l=0}^{L}B_{1,lit}\eta_{1,i,t-l}+\sum_{l=0}^{L}Q_{lit}Y_{1,i,t-l}+\sum_{l=0}^{L}\Gamma_{1,lit}X_{1,i,t-l}+\xi_{1,it} \quad (7)$$

其中，$\eta_{1,it}$表示$Y_{1,it}$对应的潜变量，Λ为因子负载矩阵，R、K、B、Q、Γ表示回归系数矩阵。$\varepsilon_{1,it}$、$\xi_{1,it}$表示残差，均值均为 0。另外, DSEM 框架在水平 2 及水平 3 的模型公式如下(Asparouhov et al., 2018)：

$$Y_{2,i}=\nu_2+\Lambda_2\eta_{2,i}+K_2X_{2,i}+\varepsilon_{2,i} \quad (8)$$

$$\eta_{2,i}=\alpha_2+B_2\eta_{2,i}+\Gamma_2X_{2,i}+\xi_{2,i} \quad (9)$$

$$Y_{3,t}=\nu_3+\Lambda_3\eta_{3,t}+K_3X_{3,t}+\varepsilon_{3,t} \quad (10)$$

$$\eta_{3,t}=\alpha_3+B_3\eta_{3,t}+\Gamma_3X_{3,t}+\xi_{3,t} \quad (11)$$

其中，$X_{2,i}$为一组个体特异(在不同个体间存在差异，但不随时间变化)的变量，$X_{3,t}$为一组时间特异(在不同时间节点间存在差异，但不随个体变化)的变量。$\eta_{2,i}$和$\eta_{3,t}$分别为$Y_{2,i}$和$Y_{3,t}$对应的潜变量。总的来说, DSEM 在水平 1 考虑了时间节点t的测量与当前及先前L次测量的回归关系，在水平 2 考虑个体特异的效应；而且，与传统多层模型不同的是, DSEM 还在水平 3 考虑了时间特异的效应。

具体来说, DSEM 在水平 1 中的回归系数、因子负载和截距项等参数的随机效应可以进一步在水平 2 和水平 3 进行解释，同时考虑不同个体和时间节点的特异贡献(Asparouhov et al., 2018)。对于水平 1 的系数s(如，s可以指代回归系数矩阵R_{lit}中的元素、因子负载矩阵$\Lambda_{1,lit}$中的元素、截距项$\alpha_{1,it}$等)，其随机效应可以在水平 2 和水平 3 被分解为$s_{2,i}$和$s_{3,t}$两个成分，并分别被个体特异和时间特异的变量解释(Asparouhov et al., 2018)。如，在最完整的 DSEM 框架下，个体积极情绪的自回归系数的随机效应将被分解为个体和时间两个方面的来源。这种随机效应在水平 2 可能受到人格特质(个体特异的变量)的影响，同时在水平 3 受到特定工作日(时间特异的变量)的影响。类似地，对于模型中的残差，其方差的随机效应也可以抽象地表示为指数函数的形式$V=\mathrm{Exp}(s_{2,i}+s_{3,t})$，从而可以分别利用个体和时间特异的变量对其随机效应进行解释(Asparouhov et al., 2018)。

McNeish 和 Hamaker (2019)指出，相比于 MLM，DSEM 的亮点在于它可以很方便地考虑变量的自回归效应和交叉滞后的效应。标准的 MLM 软件也可以创造滞后时间点的新变量加入模型，从而进行和 DSEM 类似的分析；但当时间间隔不等或存在缺失数据时, MLM 中的缺失数据点也随之增多，从而对参数估计带来更严重的问题(McNeish & Hamaker, 2019)。而 DSEM 是利用卡尔曼滤波(Kalman Filter; Harvey, 1989)，基于上一时间点的观测对当前缺失时间点的数据迭代地进行预测和更新，不需要依赖按列删除(Listwise Deletion)，因此在处理缺失数据方面有较大的优势。另外，如 3.2 节所述，在 MLM 中，研究者往往需要通过中心化来分解组间水平和组内水平的效应(Curran & Bauer, 2011)，但传统的显变量中心化在密集追踪数据中表现不佳(Asparouhov & Muthén, 2019a)。DSEM 允许对潜变量进行建模，采用潜变量中心化的方法(即在结构方程模型部分，自变量X被分为$X_{1,it}$、$X_{2,i}$、$X_{3,t}$三个成分分别进行建模)来排除来源于个体和时间的一般性影响(Asparouhov et al., 2018)。在大多数情况下，这种潜变量中心化的方法在参数估计方面的表现均比传统显变量中心化的方法更优，能很好地解决传统的显变量中心化带来的问题(Asparouhov & Muthén, 2019a)。

需要注意的是, DSEM 要求时间节点t为整数，但在实际数据中这一条件很可能无法满足(Asparouhov et al., 2018)。例如，各测量节点之间的时间间隔可能参差不齐(West, 2018)，使得时间节点t不为整数。为了解决这一问题, DSEM 需要建立另一个度量网络，对时间节点进行转换和处理，从而得到一系列近似的整数值(Asparouhov et al., 2018)。例如，当四个测量节点的间隔分别为 0.5 天、0.9 天和 2.1 天，可以选择一个较小的时间间隔$\delta=0.5$作为网格的最小单位，将四个测量时间节点近似转化为$t_0=0$, $t_1=1$, $t_2=2$, $t_4=4$。由于t_3不存在观测值, DSEM 将该节点视为缺失值处理(Asparouhov et al., 2018)。Asparouhov 等人(2018)指出，过小的时间间隔δ会导致较多的缺失值，而间隔偏大则可能会导致多个时间节点转化为同一个整数，因此需要慎重考虑δ的设置。对于如

何选择较为合适的时间间隔 δ 值，Asparouhov 等人(2018)提供了较为详细的建议。

一般来说，贝叶斯估计不需要依赖渐进理论，在小样本的情况下相比频率学派的方法具有潜在的优势(van de Schoot et al., 2014)，但这种优势不是仅通过单纯改变估计方法便能得以实现的(Depaoli & Clifton, 2015; van de Schoot et al., 2015)，往往还需要研究者结合实际情况对参数设置合适的先验分布。与其他许多贝叶斯框架下的模型类似(Depaoli & van de Schoot, 2017; McNeish, 2017; Zondervan-Zwijnenburg et al., 2017)，在样本量较小时使用 M*plus* 默认的无信息先验会使得 DSEM 更倾向于产生有偏的估计结果(McNeish, 2019)。因此，在样本量较小时，建议研究者尽可能结合前人研究和其他先验知识对参数设置有信息的先验分布；在缺乏先验信息的情况下，即使是根据参数取值范围的直观判断来对参数设置容许范围限制的先验分布(Admissible-Range-Restricted Priors)，也可以较大程度上改善 DSEM 的参数估计结果(McNeish, 2019)。

4.1.2 模型拟合及比较

DSEM 可以利用偏差信息准则(Deviance Information Criterion, DIC)作为模型拟合与比较的指标(Asparouhov et al., 2018)。DIC 常用于贝叶斯估计中的模型比较(Spiegelhalter et al., 2002)，DIC 越小的模型拟合越好。然而，当 DSEM 中参数数量较多、模型过于复杂时，需要较多迭代次数才能得到较为稳定的 DIC (Asparouhov et al., 2018)。因此，Asparouhov 等人(2018)建议，研究者需要通过对模型设置不同的随机种子并多次运行模型，以确认 DIC 是否已经达到稳定的取值。

4.1.3 优势及不足

DSEM 在密集追踪数据的分析上存在诸多亮点和优势。首先，DSEM 结合多种模型方法的优势，能更灵活、方便地对模型中多个变量之间建立相关关系(Hamaker et al., 2018; Zhou et al., 2020)。它不仅局限于对两变量的建模分析，也可以对三个及以上变量进行建模。不管是显变量还是潜变量、是否是在同一时间节点测量，DSEM 中各个变量之间均可建立回归路径(Asparouhov et al., 2018)。其次，DSEM 能方便地将密集追踪数据中呈现的特征纳入模型(如，考虑数据中存在的动态趋势、周期变化等)，不仅可以考虑变量本身随时间的变化，还可以考虑变量间关系随时间的动态变化(Asparouhov et al., 2018; Zhou et al., 2020)。另外，DSEM 对模型参数的估计采用贝叶斯模型的框架，相比于传统的频率学派方法，贝叶斯方法对于复杂模型的处理更有优势(张沥今 等, 2019; Zhou et al., 2020)。Song 和 Zhang (2014)曾将多层结构方程模型与 DFM 相结合，提出多层的动态因子模型(Multilevel Dynamic Factor Model, MDFM)，但当 MDFM 中存在随机效应时，往往会导致模型无法估计的问题。而 DSEM 对于测量节点和随机效应的设置没有数量的限制，即使存在较多的测量时间节点和随机效应，也不会对其带来较多计算量上的负担(Asparouhov et al., 2018)。前人模拟研究表明，DSEM 在随机效应的均值估计方面表现十分优秀(Schultzberg & Muthén, 2018)。

然而，DSEM 也存在一些不足与局限。在对于间隔不等的时间节点的处理中，不同时间间隔 δ 的选择会对 DSEM 水平 1 参数的估计结果产生一定的影响(Asparouhov et al., 2018; Asparouhov & Muthen, 2019b)。Asparouhov 等人(2018)在 DSEM 的基础上做出微调，提出了残差动态结构方程模型(Residual Dynamic Structural Equation Model, RDSEM)，仅在测量模型中体现自回归和交叉滞后的效应，因此其结构方程模型部分不会受到不同时间间隔选择的影响(Asparouhov et al., 2018)。Asparouhov 和 Muthén (2019b)的模拟研究表明，在观测时间节点的间隔不一的情况下，RDSEM 的模型结果相对 DSEM 来说更加稳健，几乎不会受到时间间隔的影响。

总的来说，DSEM 在密集追踪数据的分析上存在诸多亮点和优势(Hamaker et al., 2018)，但也仍面临许多新的挑战，亟待更多心理学、统计学等领域的研究者共同解决。

4.2 自下而上的方法：组迭代多模型估计

本质上来说，DSEM 和传统 MLM 一样，将被试视为整体的、同质化的群体，认为在个体间存在相同的模型路径，不同被试之间可能只是效应大小有所不同(这些随机效应可以利用个体特异和时间特异的变量进行解释)，因此是一种基于整体的方法(Piccirillo & Rodebaugh, 2019)。然而，基于整体和基于个人的分析方法可能得到截然不同的结果(Fisher et al., 2018; Turner & Hayes, 1996)，特别是实际中个体特异性较强的情况下，针对整

体数据构建的群体模型可能不能解释群体内部的任一个体(Miller & van Horn, 2007; Molenaar & Campbell, 2009)。即使患有同一疾病，不同个体之间的具体情况、变化模式等也存在较大的个体差异(Hayes et al., 2019)。因此，近年来越来越多学者倡导研究者开展基于个体的研究和统计分析(Foster & Beltz, 2018; Hayes et al., 2019; Piccirillo & Rodebaugh, 2019; Wright & Woods, 2020)。利用基于个人的分析方法对不同时间、不同情境下的动态变化过程进行建模和分析，可以得到更精确的、更具针对性的结果，从而增加治疗和干预对于个人的匹配度和治疗效用(Hayes et al., 2019; Wright & Woods, 2020)。

纯粹基于个人的分析方法充分考虑了个体间的异质性，但同时也受到了一些质疑。如，仅仅基于个人建立的模型可能不够可靠，容易受到数据中噪音的影响，从而得到过拟合的模型(Gates & Molenaar, 2012)，缺乏一定的普适性(Spencer & Schöner, 2003)。由此, Gates 和 Molenaar (2012)在针对个人的模型的基础上，结合群体共享信息的优势，提出 GIMME 这一算法。GIMME 最初是作为一种针对功能性磁共振成像(Functional Magnetic Resonance Imaging, fMRI)的数据分析方法提出，但近年来也被较多地应用于行为和心理构念方面密集追踪数据的分析(Foster & Beltz, 2018; Gates et al., 2020)，它可以被视作纯粹基于个体的方法和基于整体的方法之间的一种连接(Beltz et al., 2016; Wright & Woods, 2020)。

4.2.1 方法原理

GIMME 是一种动态网络的方法，每个变量均被视为网络中的一个节点。对于一组时间序列观测变量 $\eta_i(t)$ $(p\times 1)$，GIMME 的公式如下所示(Beltz & Gates, 2017)：

$$\eta_i(t)=(A_i+A_i^g)\eta_i(t)+(\Phi_{1,i}+\Phi_{1,i}^g)\eta_i(t-1)+\xi_i(t) \quad (12)$$

其中，A $(p\times p)$表示相同时间点各个变量节点之间的边(Contemporaneous Edges)，可以理解为变量间的同时效应。Φ $(p\times p)$为各变量节点间一阶的滞后效应(First Order Lagged Edges)，表示在时间点 t 时，变量节点与上一时间点的测量之间的关系，可能同时包括自回归和交叉滞后的效应。$\xi_i(t)$ $(p\times 1)$表示均值为 0 的误差项。GIMME 的个体模型是基于统一结构方程模型(Unified Structural Equation Modeling, USEM; Kim et al., 2007)建立。和 DFM 不同, USEM 是针对显变量(而不是潜变量)进行建模(Kim et al., 2007)。在 USEM 的基础上，GIMME 还区分了个体和群体的效应，下标 i 表示该系数为第 i 个被试独有，而上标 g 表示群体模型中的效应。因此，在各个个体的模型中，将同时包括个体特异以及群体共享的路径。简单来说, GIMME 的建模分析步骤大致包括：(1)数据预处理；(2)基于大多数个体模型中存在的路径，自下而上地建立群体的模型；(3)以群体模型为初始模型建立个体的模型(Beltz & Gates, 2017; Gates & Molenaar, 2012)。

GIMME 假设认为，残差项 $\xi_i(t)$ 与时间相互独立，不考虑变量随时间发展的线性趋势(Beltz & Gates, 2017)。因此，研究者需要将原始数据对时间这一变量进行回归分析，去除时间对变量的线性影响后，再将得到的残差项纳入 GIMME 中进行正式的分析(Beltz & Gates, 2017)。另外，研究者也可以将时间作为外源变量(Exogenous Variable)纳入模型，从而去除各变量随时间变化的趋势(Woods et al., 2020)。

在群体模型的建立中, GIMME 将初始模型作为起点，根据修正指数(Modification Indices)对个体模型逐步迭代地添加回归路径，只有显著优化大多数个体模型拟合的路径才能被纳入群体模型中(Gates & Molenaar, 2012)。接着，GIMME 以群体模型作为迭代的初始模型，对各个个体模型逐步添加显著优化其模型拟合的路径，得到各个个体的模型。关于 GIMME 建立群体和个体模型的具体算法可见 Gates 和 Molenaar (2012)的文章。

一般来说，GIMME 的一阶延迟模型(包括同时效应和滞后一个时间点的效应)已经能较好地解释密集追踪数据中的时间依赖性，但还是有必要对每个个体的模型进行后验模型验证(A Posteriori Model Validation)，以查看是否需要考虑更高阶的模型(Beltz & Gates, 2017; Beltz & Molenaar, 2015)。这种后验模型的验证可以通过检查残差项 $\xi_i(t)$ 是否为白噪音来实现(Beltz & Gates, 2017)。Beltz 和 Molenaar (2015)详细介绍了 GIMME 中后验模型验证的方法和原理，并提供了相应的 LISREL 代码。

在 GIMME 的基础上，后续较多研究也提出了其拓展模型，如 Gates 等人(2017)提出的考虑子群组的 GIMME (Subgrouping Group Iterative Multiple

Model Estimation, S-GIMME), Beltz 和 Molenaar (2016)提出的多解决方案的 GIMME (GIMME for Multiple Solutions, GIMME-MS), 以及 Gates 等人(2020)提出的基于潜变量进行建模的 GIMME (Latent Variable GIMME, LV-GIMME)。在 GIMME 的基础上, LV-GIMME 利用带有工具变量(Instrumental Variables)的两阶段最小二乘法估计(Two-Stage Least Squares Estimation, MIIV-2SLS), 同时允许测量模型和结构模型部分在不同个体间存在差异。模拟研究的结果表明, 相比于对条目总分直接进行建模, LV-GIMME 能更有效地检测出模型中存在的效应(Gates et al., 2020)。

4.2.2 优势及不足

总的来说, GIMME 在密集追踪数据的分析上具有其独特的优势。与基于整体的分析方法相比, GIMME 的优势主要体现在: (1)对于群体模型的建立是一个自下而上的过程, 其路径的纳入与否取决于该路径能否显著改善一定比例的单个个体模型(Gates & Molenaar, 2012), 可以尽可能减少个体特异信息的损失, 更适用于探究变量间共变模式的个体差异(Beltz et al., 2016); (2)允许各个个体的模型结构之间存在异质性, 并且能提供任一个体相对于整个群体的差异方面的信息, 从而更好地理解个体模型的特征和模式, 这对于临床上针对个人的精准治疗来说具有重要的意义(Wright & Woods, 2020); (3)对于被试数量的要求较少, 在只有 10 名被试的情况下就已经可以较好地对群体模型的路径进行估计(Gates & Molenaar, 2012; Henry et al., 2019)。而与纯粹基于个体的方法相比, GIMME 也具有其独特的优势: 它在考虑个体特异性的同时大大提高了结果的普适性, 一定程度上降低了模型过拟合的风险(Beltz & Gates, 2017)。另外, GIMME 在软件操作上较为简便, 且能自动输出每个个体模型以及群体模型的各项拟合指标、路径系数等, 为研究者基于个人的精细化研究提供了丰富的信息。GIMME 还可以较为直观地对结果进行呈现, 更易于理解与解释, 适用于动态系统中存在较多相互影响的变量的情况(Beltz et al., 2016)。GIMME 不仅可以帮助研究者从宏观的视角理解变量间的关系, 还可以输出每个个体的模型参数估计结果, 大大增加治疗效用和干预指导的精确性。

虽然 GIMME 在对密集追踪数据的分析上存在诸多优势, 但它同样存在一些局限。首先, GIMME 本质上是一种数据驱动的建模方法, 可能出现模型过拟合(Overfitting)等问题(Beltz & Gates, 2017)。其次, GIMME 主要考虑变量间共变(Covary)的关系及机制, 无法将几乎不随时间变化的变量纳入模型进行分析(Beltz & Gates, 2017)。为了探究个体水平上不随时间变化的协变量的作用, 研究者需要额外对个体模型中得到的参数估计进行进一步分析。如, Kim 等人(2007)利用一般线性模型(General Linear Model), 探究个体层面的变量(性别、年龄、智商及教育)对 USEM 中得到的个体模型路径的影响。又如, Beltz 等人(2013)利用 *t* 检验, 比较 USEM 得到的个体模型路径中, 各变量间达到显著的回归路径的数量是否存在显著的性别差异。另外, GIMME 假设认为变量间的效应相对时间静止, 不考虑数据中可能存在的时变性质(Time-Varying Nature; Beltz & Gates, 2017; Foster & Beltz, 2018), 但实际研究中可能出现变量间关系随时间动态变化的情况。最后, GIMME 要求各个测量时间节点之间必须间隔相等, 但在许多密集追踪的研究中, 观测节点的时间间隔常常是不等的(Foster & Beltz, 2018; Kleiman et al., 2019)。另外, 虽然 GIMME 对于个体数量上几乎没有限制, 但它要求对于每个个体进行较多时间节点的测量。

4.3 实证演示及方法比较

DSEM 已经可以在 M*plus* 上实现, 且代码语句较为简单, 这为研究者们利用 DSEM 进行密集追踪数据的动态分析提供了很大的便利(McNeish & Hamaker, 2019), 近年来涌现出较多关于 DSEM 的应用研究(Armstrong et al., 2019; Hamaker et al., 2018; Racine et al., 2019; Zhang & Zheng, 2019)。而 GIMME 已有与其同名的 R 包(*gimme*; Lane et al., 2020), 也可以利用 LISREL 以及 MATLAB 中的工具箱进行建模, 近年来在应用领域也出现较多研究利用 GIMME (Beltz & Gates, 2017; Bouwmans et al., 2018)或 GIMME 的扩展模型进行密集追踪数据的分析, 如考虑子群体模式的 S-GIMME (Lane et al., 2019), 以及可以同时输出多个群体模型方案的 GIMME-MS (Beltz et al., 2016)。

为了进一步加深读者对于 DSEM 和 GIMME 实际建模过程的理解、突出二者相比传统 MLM

的建模优势，本节将利用 M*plus* 8.4 和 R 包 *gimme*，对 McNeish 和 Hamaker (2019)提供的公开数据分别进行 MLM、DSEM 和 GIMME 的建模分析(具体代码见网络版附录)。简单来说，McNeish 和 Hamaker (2019)对 100 名被试在 50 个时间点的抑郁水平(Depression)和吸烟欲望(Urge to Smoke)进行密集追踪，利用 DSEM 一系列不同形式的模型探讨抑郁水平和吸烟欲望之间的动态关系，以及个体特异的变量(工作压力和家庭压力)在其中所起的作用。

4.3.1　传统多层线性模型的实证演示

McNeish 和 Hamaker (2019)的数据中同时包括个体层面和时间点层面的变量测量，且因变量吸烟欲望的组内相关系数(Intraclass Correlation, ICC)较大(ICC = 0.387)，因此可以利用 MLM 进行数据分析。假设吸烟欲望的基线水平、随时间变化的趋势以及抑郁对吸烟欲望的影响在不同个体间存在较大的差异，因此本例将建立带随机截距和随机斜率的 MLM 模型(对应 M*plus* 代码见网络版附录 A)。为了更好地解释结果，我们对水平 1 的预测变量(抑郁水平)进行个体均值中心化。根据 3.1 节中对于 MLM 方法原理的介绍，在水平 1，个体 i 在测量时间点 t 的吸烟欲望 Y_{it} 可以表示为：

$$Y_{it} = \beta_{0i} + \beta_{1i}(X_{it} - \bar{X}_i) + \beta_{2i}T_{it} + \varepsilon_{it} \quad (13)$$

其中，X_{it} 和 $\bar{X}_i$ 分别表示个体 i 在第 t 个测量时间点的抑郁水平以及个体 i 的平均抑郁水平，T_{it} 表示测量时间。我们假设水平 1 各项斜率和截距在不同个体间存在变异，且这一变异可能可以被个体本身来自工作和家庭的压力解释。因此，在水平 2, MLM 的模型表达式如下：

$$\beta_{0i} = \gamma_{00} + \gamma_{10}W_{工作,i} + \gamma_{20}W_{家庭,i} + u_{0i} \quad (14)$$

$$\beta_{1i} = \gamma_{01} + \gamma_{11}W_{工作,i} + \gamma_{21}W_{家庭,i} + u_{1i} \quad (15)$$

$$\beta_{2i} = \gamma_{02} + \gamma_{12}W_{工作,i} + \gamma_{22}W_{家庭,i} + u_{2i} \quad (16)$$

其中，$W_{工作,i}$ 和 $W_{家庭,i}$ 均为水平 2 的预测变量，分别表示个体的工作压力和家庭压力。MLM 对应的模型图如图 1 所示。

M*plus* 8.4 默认对 MLM 采用稳健极大似然估计(Robust Maximum Likelihood Estimator, MLR)，因此仅允许输出非标准化的结果(Muthén & Muthén, 1998-2017)。如果需要输出标准化的结果，可以采用贝叶斯估计的方法。结果显示，工作压力和家庭压力越高的个体往往有更强烈的吸烟欲望(γ_{10} = 0.458, γ_{20} = 0.286, ps < 0.001)。个体的吸烟欲望随时间的推移而逐渐增加(γ_{02} = 0.157, p < 0.001)，但这种随时间变化的趋势在个体间的随机效应不能被个体本身的工作压力或家庭压力解释(ps > 0.05)。抑郁水平越高时，个体的吸烟欲望也越强烈(γ_{01} = 0.741, p < 0.001)，且个体的工作压力和家庭压力越大，该个体抑郁水平对吸烟欲望的影响也越大(γ_{11} = 0.270, p = 0.001; γ_{21} = 0.334, p < 0.001)。总之，个体本身的工作压力和家庭压力会提高个体吸烟欲望的基线水平，且增加其抑郁水平对吸烟欲望的影响。

4.3.2　动态结构方程模型的实证演示

与 MLM 相比, DSEM 可以方便地对密集追踪数据中的动态效应进行建模分析。本例仅考虑了个体特异的预测变量，因此采用两层的 DSEM。McNeish 和 Hamaker (2019)的数据中，吸烟欲望整体呈现出随时间线性增长的趋势，因此需要对数据进行去时间趋势(Detrending)。Asparouhov

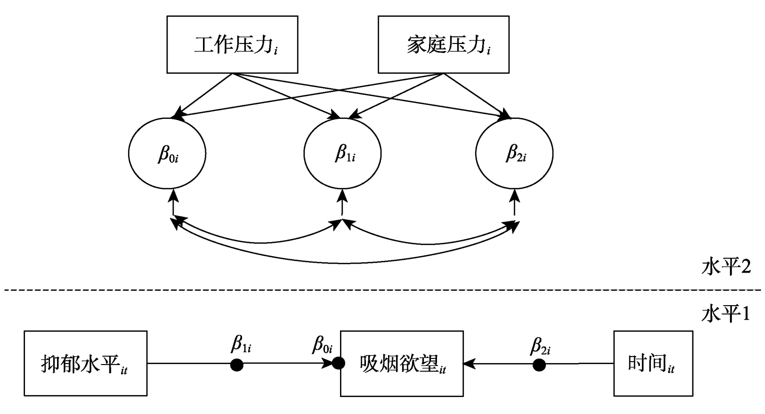


图 1　抑郁水平对吸烟欲望的影响

注：圆圈内变量为潜变量，方框内变量为显变量。下标 i 表示第 i 个个体的测量，t 表示第 t 个时间点的观测数据。

(2018)的模拟研究表明, DSEM 在同时进行去时间趋势和自回归效应的分析时会对方差的估计产生偏差, 因此建议采用DSEM的变式——RDSEM进行分析。RDSEM 与 DSEM 的区别主要在于RDSEM 分别对去时间趋势与自回归和交叉滞后效应进行建模, 即仅在测量模型部分而不是结构方程模型的部分考虑自回归和交叉滞后效应。RDSEM 是对 DSEM 微调后的提出的一种变式(Asparouhov et al., 2018), 因此对 RDSEM 的介绍同样可以较好地展示 DSEM 的分析过程及思路。

根据理论假设, 在水平 1, 吸烟欲望和抑郁水平存在自回归($\Phi_{\text{吸烟–吸烟}}, \Phi_{\text{抑郁–抑郁}}$)和交叉滞后的效应, 即前一个时间点的抑郁水平会对后一时间点的吸烟欲望产生影响($\Phi_{\text{吸烟–抑郁}}$), 而前一时间点的吸烟欲望也可能会对后一时间点的抑郁水平产生作用($\Phi_{\text{抑郁–吸烟}}$)。对应模型公式为:

$$\begin{bmatrix} Y_{\text{吸烟},it} \\ Y_{\text{抑郁},it} \end{bmatrix} = \begin{bmatrix} \mu_{\text{吸烟},i} \\ \mu_{\text{抑郁},i} \end{bmatrix} + \begin{bmatrix} Y^{(w)}_{\text{吸烟},it} \\ Y^{(w)}_{\text{抑郁},it} \end{bmatrix} \quad (17)$$

$$\begin{bmatrix} Y^{(w)}_{\text{吸烟},it} \\ Y^{(w)}_{\text{抑郁},it} \end{bmatrix} = \begin{bmatrix} \beta_{\text{吸烟–时间},i} \\ \beta_{\text{抑郁–时间},i} \end{bmatrix} T_{it} + \begin{bmatrix} \varsigma_{\text{吸烟},it} \\ \varsigma_{\text{抑郁},it} \end{bmatrix} \quad (18)$$

$$\begin{bmatrix} \varsigma_{\text{吸烟},it} \\ \varsigma_{\text{抑郁},it} \end{bmatrix} = \begin{bmatrix} \Phi_{\text{吸烟–吸烟},i} & \Phi_{\text{吸烟–抑郁},i} \\ \Phi_{\text{抑郁–吸烟},i} & \Phi_{\text{抑郁–抑郁},i} \end{bmatrix} \begin{bmatrix} \varsigma_{\text{吸烟},i,t-1} \\ \varsigma_{\text{抑郁},i,t-1} \end{bmatrix} + \begin{bmatrix} \delta_{\text{吸烟},it} \\ \delta_{\text{抑郁},it} \end{bmatrix} \quad (19)$$

与 4.1.1 节中 DSEM 的公式相比, RDSEM 仅在测量模型部分考虑残差项 ς 的自回归和交叉滞后的效应。令 $\pi_{\text{吸烟},i} = \text{var}(\delta_{\text{吸烟},it}), \pi_{\text{抑郁},i} = \text{var}(\delta_{\text{抑郁},it})$, 则在水平 2, 个体的吸烟欲望和抑郁的平均水平($\mu_{\text{吸烟},i}, \mu_{\text{抑郁},i}$)、各个回归系数($\Phi_{\text{吸烟–吸烟},i}, \Phi_{\text{抑郁–抑郁},i}, \Phi_{\text{抑郁–吸烟},i}, \Phi_{\text{吸烟–抑郁},i}, \beta_{\text{吸烟–时间},i}, \beta_{\text{抑郁–时间},i}$), 以及残差项方差($\pi_{\text{吸烟},i}, \pi_{\text{抑郁},i}$)均可能受到个体的工作压力和家庭压力的影响。在考虑残差项方差的随机效应时, 为了保证残差项方差为正数, DSEM 将取对数之后的残差项方差 $\pi_{\text{吸烟},i}$ 和 $\pi_{\text{抑郁},i}$ 纳入水平 2 的模型。因此, 水平 2 模型的公式为:

$$\begin{bmatrix} \mu_{\text{吸烟},i} \\ \mu_{\text{抑郁},i} \\ \Phi_{\text{吸烟–吸烟},i} \\ \Phi_{\text{吸烟–抑郁},i} \\ \Phi_{\text{抑郁–吸烟},i} \\ \Phi_{\text{抑郁–抑郁},i} \\ \beta_{\text{吸烟–时间},i} \\ \beta_{\text{抑郁–时间},i} \\ \log(\pi_{\text{吸烟},i}) \\ \log(\pi_{\text{抑郁},i}) \end{bmatrix} = \begin{bmatrix} \gamma_{\text{吸烟}} \\ \gamma_{\text{抑郁}} \\ \gamma_{\text{吸烟–吸烟}} \\ \gamma_{\text{吸烟–抑郁}} \\ \gamma_{\text{抑郁–吸烟}} \\ \gamma_{\text{抑郁–抑郁}} \\ \gamma_{\text{吸烟–时间}} \\ \gamma_{\text{抑郁–时间}} \\ \gamma_{\log(\pi_{\text{吸烟}})} \\ \gamma_{\log(\pi_{\text{抑郁}})} \end{bmatrix} + W_{\text{工作},i} \begin{bmatrix} \gamma_{\text{工作, 吸烟}} \\ \gamma_{\text{工作, 抑郁}} \\ \gamma_{\text{工作, 吸烟–吸烟}} \\ \gamma_{\text{工作, 吸烟–抑郁}} \\ \gamma_{\text{工作, 抑郁–吸烟}} \\ \gamma_{\text{工作, 抑郁–抑郁}} \\ \gamma_{\text{工作, 吸烟–时间}} \\ \gamma_{\text{工作, 抑郁–时间}} \\ \gamma_{\text{工作}, \log(\pi_{\text{吸烟}})} \\ \gamma_{\text{工作}, \log(\pi_{\text{抑郁}})} \end{bmatrix} + W_{\text{家庭},i} \begin{bmatrix} \gamma_{\text{家庭, 吸烟}} \\ \gamma_{\text{家庭, 抑郁}} \\ \gamma_{\text{家庭, 吸烟–吸烟}} \\ \gamma_{\text{家庭, 吸烟–抑郁}} \\ \gamma_{\text{家庭, 抑郁–吸烟}} \\ \gamma_{\text{家庭, 抑郁–抑郁}} \\ \gamma_{\text{家庭, 吸烟–时间}} \\ \gamma_{\text{家庭, 抑郁–时间}} \\ \gamma_{\text{家庭}, \log(\pi_{\text{吸烟}})} \\ \gamma_{\text{家庭}, \log(\pi_{\text{抑郁}})} \end{bmatrix} + \begin{bmatrix} u_{\text{吸烟},i} \\ u_{\text{抑郁},i} \\ u_{\text{吸烟–吸烟},i} \\ u_{\text{吸烟–抑郁},i} \\ u_{\text{抑郁–吸烟},i} \\ u_{\text{抑郁–抑郁},i} \\ u_{\text{吸烟–时间},i} \\ u_{\text{抑郁–时间},i} \\ u_{\log(\pi_{\text{吸烟},i})} \\ u_{\log(\pi_{\text{抑郁},i})} \end{bmatrix} \quad (20)$$

其中, 公式(17)对应表示变量变异来源的分解, 公式(18)~(19)对应水平 1 的模型, 公式(20)对应水平 2 的模型(对应 M*plus* 代码见网络版附录A)。McNeish 和 Hamaker (2019)的数据仅提供了个体特异预测变量的数据, 因此本例仅假设水平1 的各项系数在水平 2 受到个体特异预测变量的影响; 但如果数据信息充足, 我们也可以类似地在水平 3 考虑时间特异的预测变量(如, 特定的工作日)产生的影响, 如可以将残差项方差 $\pi_{\text{抑郁},i}$ 的随机效应抽象地表示为 $\log(\pi_{\text{抑郁},i}) = s_{2,i} + s_{3,t}$, 从而将其分解为个体和时间特异的随机效应。在本例中, 对于第 i 个个体, 其模型图如图 2 所示。

DSEM 和 RDSEM 均可以同时输出标准化和非标准化的结果, 且标准化结果还在水平 1 输出了不同个体中各项标准化系数的平均结果(Within-Level Standardized Estimates Averaged Over Clusters)。为了更好地对吸烟欲望和抑郁水平之间的交叉滞后效应进行解读, 我们采用的是 M*plus* 输出的标准化结果, 其标准化的具体过程及原理见Schuurman 等人(2016)的文章。水平 1 标准化系数在不同个体中的平均结果显示, 吸烟欲望的自回归效应达到显著($\overline{\Phi_{\text{吸烟–吸烟}}}$ = 0.177, CI = [0.142,

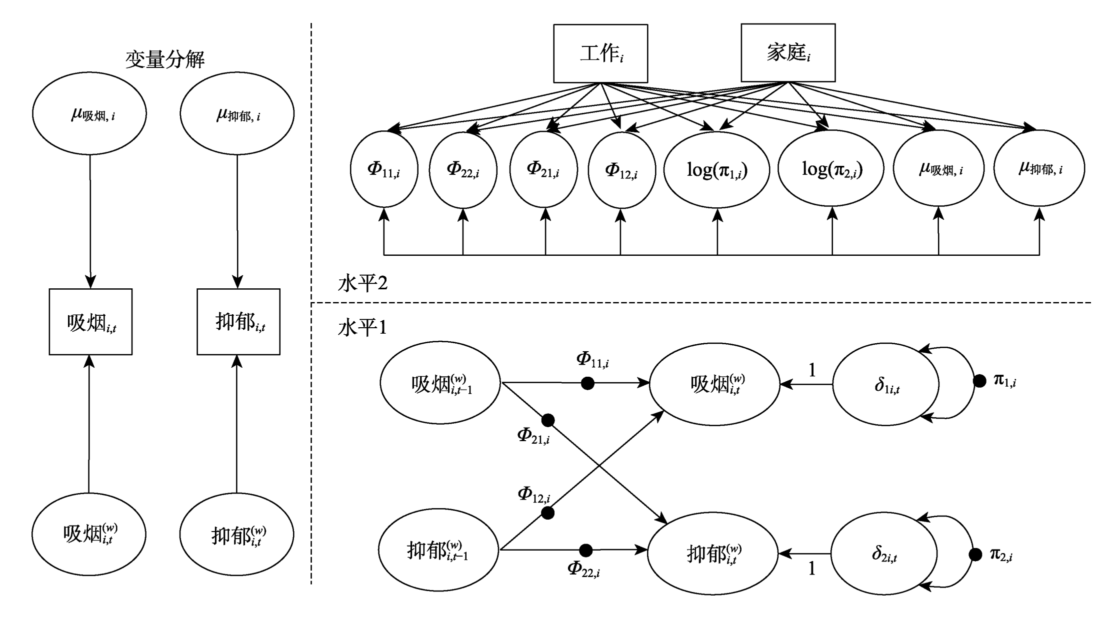


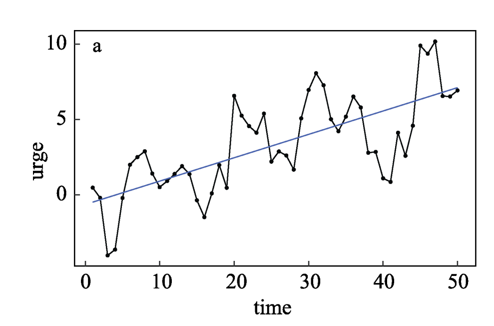
图 2　吸烟欲望和抑郁水平的动态关系

注：圆圈内变量为潜变量，方框内变量为显变量。(W)表示个体内的估计，“吸烟”和“抑郁”分别表示吸烟欲望和抑郁水平。$\Phi_{11,i}$ 和 $\Phi_{22,i}$ 分别表示吸烟欲望和抑郁水平的自回归效应，$\Phi_{21,i}$ 和 $\Phi_{12,i}$ 分别表示上一时间点的吸烟欲望对当前抑郁水平、上一时间点的抑郁水平对当前吸烟欲望的交叉滞后效应。μ 表示均值，$\delta_{吸烟,i,t}$ 和 $\delta_{抑郁,i,t}$ 分别表示吸烟欲望和抑郁水平的残差项。$\pi_{1,i}$ 和 $\pi_{2,i}$ 分别为 $\delta_{1,i,t}$ 和 $\delta_{2,i,t}$ 的方差。为了更清晰地呈现模型假设，模型图中未呈现与去时间趋势有关的路径。

0.211])，前一时间点的吸烟欲望可以显著预测后一时间点的吸烟欲望。另外，被试的吸烟欲望整体随时间的推移而逐渐增加($\overline{\Phi_{吸烟-时间}}$ = 0.721, CI = [0.704, 0.738])。前一时间点的抑郁水平对后一时间点吸烟欲望的影响没有达到显著，但前一时间点的吸烟欲望可以降低后一时间点的抑郁水平($\overline{\Phi_{抑郁-吸烟}}$ = −0.069, CI = [−0.102, −0.038])。在水平 2，工作压力和家庭压力越大的个体，其吸烟欲望越高($\gamma_{工作,吸烟}$ = 0.362, CI = [0.190, 0.515]; $\gamma_{家庭,吸烟}$ = 0.220, CI = [0.055, 0.371])，吸烟欲望的自回归效应越强($\gamma_{工作,吸烟-吸烟}$ = 0.423, CI = [0.220, 0.594]; $\gamma_{家庭,吸烟-吸烟}$ = 0.212, CI = [0.016, 0.388])，吸烟欲望的残差项方差越大($\gamma_{工作,\log(\pi_{吸烟})}$ = 0.355, CI = [0.227, 0.474]; $\gamma_{家庭,\log(\pi_{吸烟})}$ = 0.211, CI = [0.076, 0.334])，因此其吸烟欲望的残差的动态波动幅度也越大。RDSEM 的结果进一步表明，个体的吸烟欲望存在较强的惯性(自回归效应)，且个体的工作压力和家庭压力越大，这种惯性也越强。我们在进行 MLM 的分析中发现抑郁水平对吸烟欲望存在正向预测的作用，而 RDSEM 的结果体现了相反方向的另一种可能，即前一时间点的吸烟欲望可以降低后一时间点的抑郁水平。现实生活中，变量间的相互作用很可能存在复杂的双向反馈，而 DSEM 允许在同一模型中建立双向的作用路径，这可能可以为研究者带来更多新的研究思路和启发。

4.3.3　组迭代多模型估计的实证演示

本节将利用 *gimme* 这一 R 包进行 GIMME 的建模分析。在对数据进行分析之前，首先需要去除数据中随时间线性变化的趋势。研究者可以将原始数据对时间进行回归后得到的残差项纳入 GIMME 中进行正式的分析，也可以直接将时间作为外源变量纳入模型，从而去除各变量随时间变化的趋势(Woods et al., 2020)。为了更直观地呈现去时间趋势的效果，本例采用前一种去时间趋势的方法。以第一个被试的吸烟欲望为例，图 3a 为吸烟欲望的原始测量(URGE)随时间变化的总体趋势，随着时间的推移，吸烟欲望越强烈。通过建立时间对吸烟欲望的回归模型得到的残差来去除数据中随时间线性变化的趋势后，其结果如图 3b 所示。可以看出，此时吸烟欲望在不同时间点的测量呈平稳的趋势。



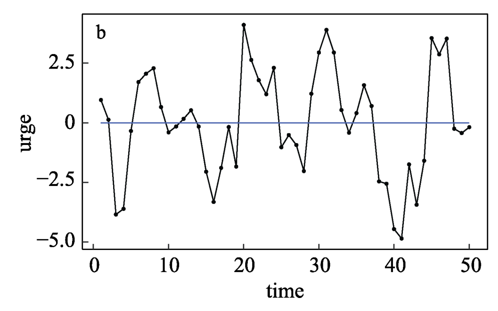


图 3 吸烟欲望随时间发生的变化
注：图 3a 为吸烟欲望的原始测量随时间变化的总体趋势；图 3b 表示去除数据中随时间线性变化的趋势之后的结果。

去除各变量受到的来自时间的影响后，将其纳入 GIMME 进行分析。对于每个被试 i，其吸烟欲望和抑郁水平的时间序列观测变量可以表示为：

$$\begin{bmatrix}\eta_{urge,i}(t)\\ \eta_{dep,i}(t)\end{bmatrix}=(A_i+A_i^g)\begin{bmatrix}\eta_{urge,i}(t)\\ \eta_{dep,i}(t)\end{bmatrix}+(\Phi_{1,i}+\Phi_{1,i}^g)\begin{bmatrix}\eta_{urge,i}(t-1)\\ \eta_{dep,i}(t-1)\end{bmatrix}+\begin{bmatrix}\xi_{urge,i}(t)\\ \xi_{dep,i}(t)\end{bmatrix} \quad (21)$$

其中，$\eta_{urge,i}(t)$ 和 $\eta_{dep,i}(t)$ 分别表示被试 i 在时间点 t 的吸烟欲望及抑郁水平。A (2×2)表示吸烟欲望和抑郁水平的同时效应矩阵，$\Phi(2\times2)$为吸烟欲望与抑郁水平之间的一阶滞后效应(First Order Lagged Edges)，包括自回归和交叉滞后效应。下标 i 表示该系数为第 i 个被试独有，而上标 g 表示为群体模型中的效应。在利用 GIMME 进行数据分析时，研究者可以利用已有的经验对群体模型中可能存在的路径做出假设，并将其作为初始迭代的模型(如，在 *gimmeSEM* 函数中声明 *paths* = *urge~deplag*，则表示初始模型中存在前一时间点的抑郁水平指向当前吸烟欲望的回归路径)。若理论依据不足，也可以直接从空模型出发，以数据为驱动进行探索式的建模(Beltz & Gates, 2017; Gates & Molenaar, 2012)。本文将利用数据驱动、探索式的方式进行演示。

在准备好数据并运行函数后，*gimme* 将初始模型作为起点，自下而上地建立群体模型。具体来说，*gimme* 首先根据修正指数对每个个体模型逐步迭代地添加回归路径，再以个体模型中的路径为基础建立群体模型。为了找到在大多数个体模型中较为稳定存在的路径和效应，该路径只有能显著优化大多数个体模型的拟合时(大多数个体模型中，该路径对应的 MI 在 0.01 的水平上达到显著)，才能被纳入群体模型中(Gates & Molenaar, 2012)。本文按照默认的设置，将 75%设为可以被纳入群体模型的临界值，研究者也可以根据具体需要进行修改。基于群体模型，*gimme* 再对个体模型逐个进行迭代路径搜索。在完成所有模型的路径搜索后，*gimme* 将输出各个模型的路径图及详细的统计结果。图 4 即 *gimme* 输出的群体模型图。群体模型结果显示，超过 75%的个体模型中，吸烟欲望和抑郁水平均存在自回归的效应($\beta_{\text{urge}\rightarrow\text{urgelag}}$ = 0.196, $\beta_{\text{dep}\rightarrow\text{deplag}}$ = −0.049)。而其他路径均没有在超过 75%的个体模型中出现，这一结果也体现了变量间关系的个体差异。

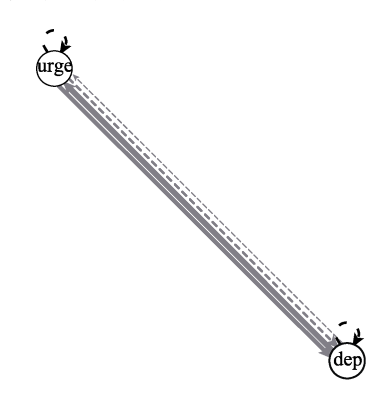


图 4 群体模型
注：urge: 吸烟欲望; dep: 抑郁水平。黑线表示群体模型中的路径，灰线表示个体模型中存在的路径，实线代表同时效应，虚线表示滞后效应(包括自回归和交叉滞后效应)。线条粗细表示模型中存在该路径的个体数量的多少。

图 5 为随机选取的个体模型示例。总体来看，不同个体在吸烟欲望和抑郁水平之间的动态关系存在较大的个体差异。如，对于个体模型中吸烟欲望的自回归效应 $\beta_{\text{urge}\rightarrow\text{urgelag}}$，较多个体在前一个时间点较高的吸烟欲望均会预测下一个时间点更高的吸烟欲望(47%)，但也有较多个体模型中

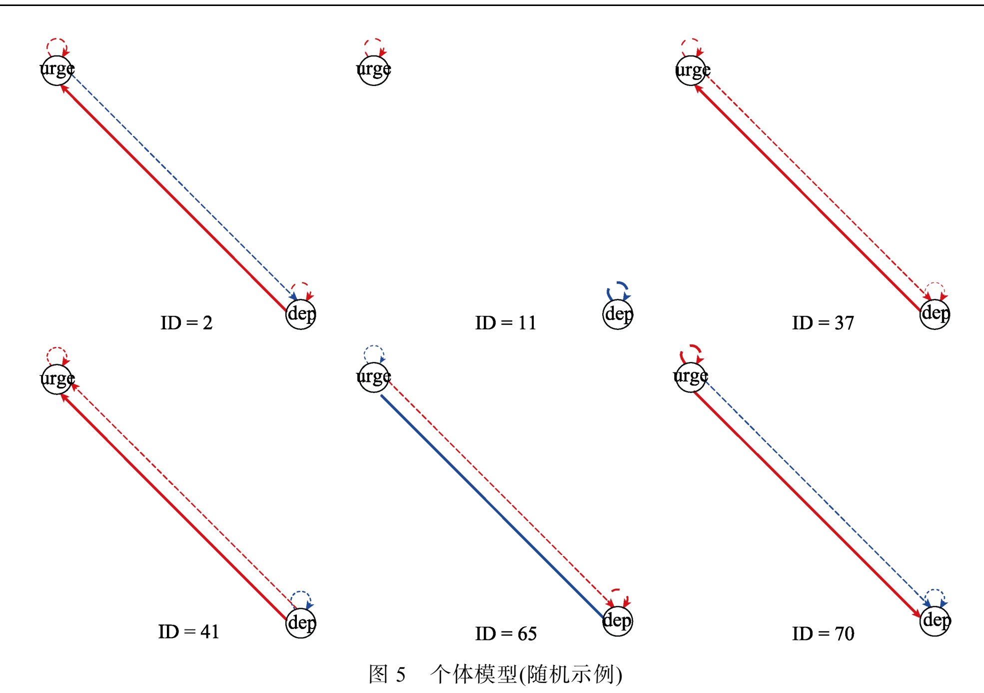


图 5　个体模型(随机示例)

注：urge: 吸烟欲望; dep: 抑郁水平。红线表示正向的效应, 蓝线表示负向的效应, 实线代表同时效应, 虚线表示滞后效应(包括自回归和交叉滞后效应)。

二者不存在预测作用(47%), 甚至可能存在负向的自回归路径(6%)。另外, 在 9 个个体模型中, 吸烟欲望可以预测下一时间点的抑郁水平, 其中 4 人为负向预测, 5 人为正向预测。

4.3.4　小结

本节分别利用 MLM、RDSEM 和 GIMME 进行实证演示和分析。RDSEM 以 DSEM 作为基础, 因此本节也较好展示了 DSEM 的分析过程及思路。MLM 和 DSEM 均为自上而下的建模方法, 相比 MLM, DSEM 在分析密集追踪数据时存在诸多优势：(1)可以方便地对变量间的自回归和交叉滞后效应进行分析(包括相反方向的交叉滞后效应); (2)可以同时考虑多个因变量; (3)对模型随机效应的数量几乎没有限制; (4)可以在水平 1 输出不同个体中各项标准化系数的平均结果; (5)采用潜变量中心化; (6)可以考虑残差项方差的随机效应。如果研究者希望更加系统、深入地学习 DSEM 在 M*plus* 中的代码实现, 可以参考 McNeish 和 Hamaker (2019)的文章。

与 MLM 和 DSEM 不同, GIMME 是一种典型的自下而上的建模方法。与 DSEM 相比, GIMME 的优势主要体现在它可以单独输出各个个体模型的结果, 允许不同个体模型的结构存在异质性, 从而更有针对性地呈现每个个体中变量相互作用的模式。然而, GIMME 不能将不随时间变化的变量纳入模型, 因此不能方便地考虑个体特异的变量(如工作压力、家庭压力)对个体内部变量间作用关系的影响。

4.4　其他模型

除了以上介绍的模型方法, 在密集追踪数据中也存在其他模型方法的应用。如, 考虑到心理学研究中常常出现潜变量的情况, 研究者们将 MLM 和 SEM 相结合, 提出多层结构方程模型(Multilevel Structural Equational Model, MSEM; Gottfredson et al., 2009; Johnson et al., 2005)。与 MLM 类似, MSEM 在探究个体内部的动态过程以及个体特异性等方面也存在较大的局限性。如研究者希望更深入地了解 MSEM, 可以参阅 Rabe-Hesketh 等(2007), Preacher 等(2010)以及 Holtmann 等(2016)的文章。

多层 VAR 模型(Multilevel Vector Autoregressive Model, mlVAR; Bringmann et al., 2015)则是在 VAR 模型的基础上, 进一步允许对多个个体的密集追踪数据进行分析, 也是目前较为常见的用于密集追踪的方法之一(Bringmann et al., 2016; Jongerling et al., 2015; Piccirillo & Rodebaugh,

2019)。特别是其与网络分析相结合, 以及其 R 包 *mlVAR* 的提出(Bringmann et al., 2013; Epskamp et al., 2016), 使得 mlVAR 的应用相对更加简便, 且同样可以输出与 GIMME 类似的模型网络, 便于在临床上识别症状网络中的症状中心(Symptom Centrality; Bringmann et al., 2015)。然而, 与 GIMME 不同, mlVAR 本质上采用的是基于整体的建模思想。mlVAR 具有测量时间点的数量要求较低、适用于高维数据分析的优点, 但它只能进行显变量的建模, 且难以对复杂模型进行估计(Bringmann et al., 2015; Epskamp et al., 2018)。由于篇幅有限, 本文仅对 mlVAR 进行简单介绍, 更详细的介绍可参见 Bringmann 等人(2015)的文章。

另外, 考虑到随着时间的推移, 变量动态变化过程中可能存在一定的时变效应(Time-Varying Effect), Bringmann 等人(2018)在 VAR 的基础上进一步将时变效应纳入考虑, 提出时变向量自回归模型(Time-Varying Vector Autoregression, TV-VAR; Bringmann et al., 2018)。TV-VAR 最初起源于计量经济学领域, 但近年来也被应用于分析心理过程的动态变化(Bringmann et al., 2017; Bringmann et al., 2018; Chow et al., 2017; Haslbeck & Waldorp, 2020)。与前面介绍的模型方法不同, TV-VAR 不需要假设时间序列的稳定性(Stationarity), 而是允许模型参数随时间发生变化(唐文清 等, 2020; Bringmann et al., 2018)。然而, TV-VAR 对于观测数据点的数量要求较高(一般需要对每个被试测量超过 100 个时间点), 且目前只能用于两变量的建模分析(Piccirillo & Rodebaugh, 2019)。感兴趣的研究者可以参阅唐文清等人(2020)以及 Bringmann 等人(2018)的文章。

4.5 密集追踪数据分析方法的选择策略

为了帮助研究者针对自身研究的实际情况, 灵活选择适合的模型方法, 这里对各种主流模型方法的优缺点及选择策略进行归类和总结(见表 1 及图 6), 以供参考(Asparouhov et al., 2018; Gates et al., 2020; Gates & Molenaar, 2012; Piccirillo & Rodebaugh, 2019):

5 讨论

密集追踪方法对个体展开多个密集时间节点的数据追踪, 能更好地研究个体的变化过程和内部机制, 在外部效度、数据分析方法等方面相比传统的追踪研究具有较大优势。为了鼓励更多的研究者在密集追踪数据中关注动态变化的过程、不同时间节点间存在的自回归效应等, 本文主要介绍了两种密集追踪模型方法的新进展, 希望能够增进研究者们对其的理解、运用和进一步发展。相比传统的分析方法, DSEM 和 GIMME 都更多地纳入了时间序列的建模思想, 并更加强调对于个体间差异的关注。DSEM 允许研究者对任意数量的随机效应进行考虑(Asparouhov et al., 2018), 而 GIMME 从纯粹的基于个人的分析方法出发, 更直接地对单个个体的模式进行探索。这些分析方

表 1 密集追踪的主流方法优缺点

模型方法	优点	缺点
MLM/MSEM	简单易行; 对时间点数量要求相对较低	假设不同个体间变量的相互作用机制是同质的; 难以分析个体内部的动态变化过程及机制, 建模不够灵活
mlVAR	对时间点数量要求相对较低; 适用于较多变量间的动态交互; 结果可视性强	假设不同个体间变量的相互作用机制是同质的; 不考虑潜变量因子结构
DSEM	可以将随机效应分解为个体和时间两个方面的来源; 在时间节点和随机效应的数量上限制较少; 允许测量时间节点之间间隔不同	假设不同个体间变量的相互作用机制是同质的; 需要对被试测量较多时间点
(LV-)GIMME	针对单个被试进行建模, 考虑个体内部的动态变化过程及机制的异质性; 适用于较多变量间的动态交互; 自动化搜索	可能出现模型过拟合; 无法将几乎不随时间变化的变量纳入模型; 需要对单个被试测量较多时间点

注: GIMME: 组迭代多模型估计(Group Iterative Multiple Model Estimation); LV-GIMME: 基于潜变量进行建模的 GIMME (Latent Variable GIMME); MLM: 多层线性模型(Multilevel Model); MSEM: 多层结构方程模型(Multilevel Structural Equational Model); mlVAR: 多层 VAR 模型(Multilevel VAR); DSEM: 动态结构方程模型(Dynamic Structural Equation Model)。下同。

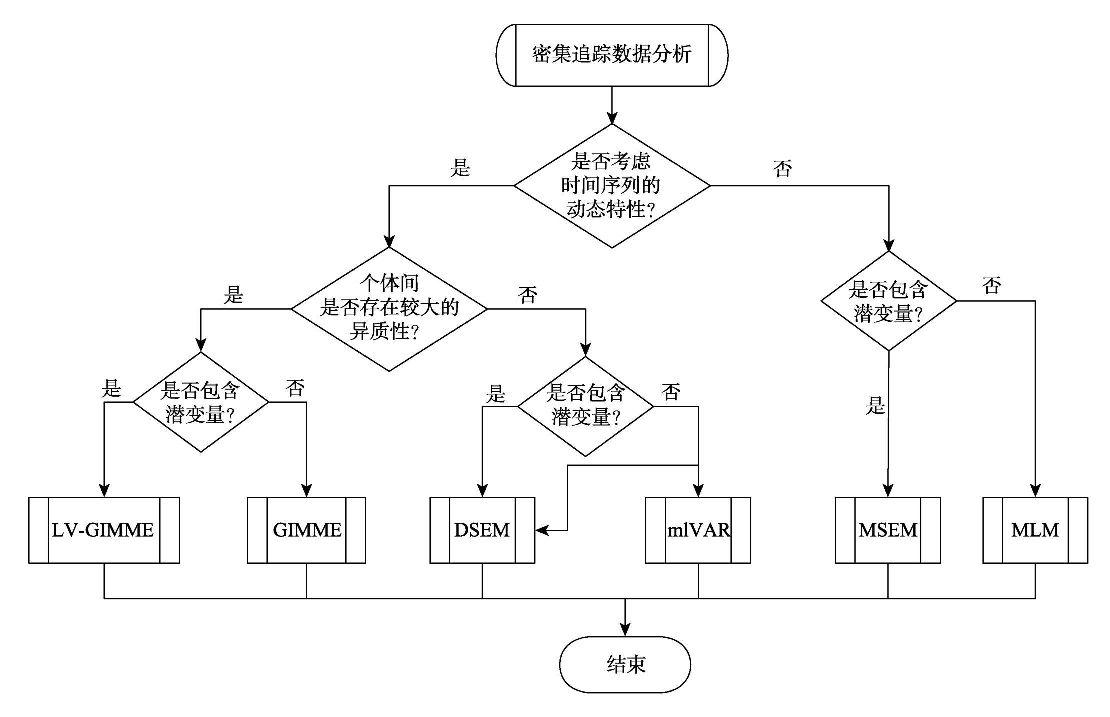


图 6　密集追踪数据方法的选择策略流程图

法对个体差异的高度关注，反映了近年来对于个体特异性研究的倡导和发展的趋势。

值得注意的是, GIMME 的一大特点在于其模型的建立是探索性或者半验证性的，这与传统心理学研究中基于前人研究及已有的知识经验提出假设的模型、利用大规模的数据收集对模型进行进一步验证的思想存在很大的不同。这种探索性的建模思路允许研究者以数据驱动的方式对变量间的关系进行自由探索，可能可以避免先入为主的固化思维，为研究者提供更多新的发现和启发。然而，探索性分析常常也伴随模型过拟合、结果难以解释等问题，未来还需要研究者进一步探究对其停止搜索规则的优化设置。DSEM 对模型中随机效应的数量几乎没有限制，这在改善模型拟合的同时也可能导致模型过于复杂、无法泛化到更广泛群体的问题。为了尽可能减少可重复性方面的问题，建议研究者在实际研究前对其实验设计及分析计划进行预注册，并在文章中尽可能详尽地报告其数据收集、处理及分析等过程。

另外，目前对于密集追踪数据的动态分析大多假设测量时间点是离散化的，这种假设在连续时间测量，以及测量时间间隔高度随机的情况下可能存在一定的问题(Hamaker et al., 2018)。连续时间建模(Continuous Time Modeling)可以较好地解决时间离散化假设带来的问题，但它需要涉及复杂的微分计算(Voelkle et al., 2018)，目前应用研究者对于如何在实证分析中实现连续时间建模还不够熟悉。未来还需要更多研究者进一步探索连续时间模型的简化计算与自动化实现，为应用研究者提供更灵活的建模选择。

密集追踪研究关注个体内行为、认知、心理等的动态变化过程，这些内在过程及机制往往十分复杂，可能存在非线性的作用模式(如，两变量间存在交互作用)。目前密集追踪数据较为主流的分析方法大多仅考虑线性效应的情况(如，DSEM 和 GIMME)，但仅通过线性关系对这些动态效应进行解释很可能是不够的。最近, Kelava 和 Brandt (2019)将 DSEM 与动态潜类别分析(Dynamic Latent Class Analysis; Asparouhov et al., 2017)相结合，提出非线性的动态潜类别结构方程模型(Nonlinear Dynamic Latent Class Structural Equation Modeling, NDLC-SEM)，允许在组间和组内水平进行非线性效应的分析(Kelava & Brandt, 2019)。然而，目前允许对密集追踪数据进行非线性建模的模型方法仍较为有限，其在不同情境下的具体表现还不明晰，未来需要更多方法学研究者对其进行更进一

步的研究。

密集追踪研究的数据收集往往基于被试的真实生活场景，具有较高的外部效度。然而，真实的测量环境往往也更加复杂，研究者可能需要相应增加对潜在干扰因素的动态测量，并将其纳入模型进行分析。此时，模型中涉及的变量数目较多，因此采用 GIMME 对数据进行分析可能是更合适的。然而, GIMME 对于单个个体的测量时间点数要求较高，这大大增加了应用研究者密集追踪数据的收集难度。未来可以进一步探索适用于高维密集追踪数据、且对测量时间点要求相对较低的模型方法，以及进一步发展密集追踪数据的收集方法，尽可能降低其收集的难度。

总之，密集追踪方法的一大优势在于探究个体内部的动态变化过程及其机制。虽然目前密集追踪数据分析在方法学领域已有较多新的进展，但它在实际研究中应用不足，许多应用研究者缺乏对于时间这一要素的关注(Cronin & Vancouver, 2019; Voelkle et al., 2018; Zhou et al., 2020)。现在，较多密集追踪数据的分析方法已经可以方便地通过软件实现(如，DSEM 和 GIMME)，大大减少了实证研究应用的阻碍。希望本文可以帮助实证研究者们了解、掌握更灵活的建模和分析方法，更好地发挥密集追踪数据的优势，尽可能减少对数据结果的误读。

参考文献

安媛媛，徐慰. (2017). 雾霾和知觉压力加重疲劳感：来自生态瞬时评估的证据. *中国临床心理学杂志, 25*(6), 1130–1133.

陈明瑞，周萍. (2017). 成瘾物质使用的生态瞬时评估与干预. *心理科学进展, 25*(2), 247–252.

任杰，金志成，何慧. (2010). 日重现法在城市居民主观幸福感测量中的应用. *心理科学, 33*(5), 1266–1269.

邵华，陈奕荣，郝思哲. (2019). 日常生活中的压力与应对: 一项生态瞬时评估研究. *中国临床心理学杂志, 27*(3), 561–565.

唐文清，张敏强，方杰. (2020). 时变效应模型及在密集追踪数据分析中的应用. *心理科学, 43*(2), 488–497.

张峰，耿晓伟. (2015). 基于日重现法的农村居民幸福感研究：情感体验的视角. *心理与行为研究, 13*(3), 391–396.

张沥今，陆嘉琦，魏夏琰，潘俊豪. (2019). 贝叶斯结构方程模型及其研究现状. *心理科学进展, 27*(11), 1812–1825.

张银普，骆南峰，石伟. (2016). 经验取样法：一种收集"真实"数据的新方法. *心理科学进展, 24*(2), 305–316.

郑文倩，潘康，陈云云，宋国萍. (2019). 基于日记法探索正念与限制性饮食：身体意象的中介作用. *心理技术与应用, 7*(12), 735–745.

Armstrong, B., Covington, L. B., Unick, G. J., & Black, M. M. (2019). Featured article: Bidirectional effects of sleep and sedentary behavior among toddlers: A dynamic multilevel modeling approach. *Journal of Pediatric Psychology, 44*(3), 275–285.

Asparouhov, T. (2018). *Dynamic structural equation modeling of intensive longitudinal data using Mplus Version 8*. Retrieved June 2, 2020, from http://www.statmodel.com/download/Part%205%20Asparouhov.pdf

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2017). Dynamic latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(2), 257–269.

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 25*(3), 359–388.

Asparouhov, T., & Muthén, B. (2019a). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling: A Multidisciplinary Journal, 26*(1), 119–142.

Asparouhov, T., & Muthén, B. (2019b). Comparison of models for the analysis of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–23.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1–39). New York, NY: Academic.

Beltz, A. M., Beekman, C., Molenaar, P. C. M., & Buss, K. A. (2013). Mapping temporal dynamics in social interactions with unified structural equation modeling: A description and demonstration revealing time-dependent sex differences in play behavior. *Applied Developmental Science, 17*(3), 152–168.

Beltz, A. M., & Gates, K. M. (2017). Network mapping with GIMME. *Multivariate Behavioral Research, 52*(6), 789–804.

Beltz, A. M., & Molenaar, P. C. M. (2015). A posteriori model validation for the temporal order of directed functional connectivity maps. *Frontiers in Neuroscience, 9*, 304.

Beltz, A. M., & Molenaar, P. C. M. (2016). Dealing with multiple solutions in structural vector autoregressive models. *Multivariate Behavioral Research, 51*(2-3), 357–373.

Beltz, A. M., Wright, A. G. C., Sprague, B. N., & Molenaar, P. C. M. (2016). Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment,*

23(4), 447–458.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*(1), 579–616.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research.* (pp. 1–9). New York: The Guilford Press

Bouwmans, M. E. J., Beltz, A. M., Bos, E. H., Oldehinkel, A. J., de Jonge, P., & Molenaar, P. C. M. (2018). The person-specific interplay of melatonin, affect, and fatigue in the context of sleep and depression. *Personality and Individual Differences, 123*, 163–170.

Bringmann, L. F., Ferrer, E., Hamaker, E. L., Borsboom, D., & Tuerlinckx, F. (2018). Modeling nonstationary emotion dynamics in dyads using a time-varying vector-autoregressive model. *Multivariate Behavioral Research, 53*(3), 293–314.

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods, 22*(3), 409–425.

Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Methods, 45*(4), 747–757.

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., & Kuppens, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment, 23*, 425–435.

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., … Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *Plos One, 8*(4), e60188.

Chen, Y., & Zhang, S. (2020). A Latent Gaussian process model for analysing intensive longitudinal data. *British Journal of Mathematical and Statistical Psychology, 73*(2), 237–260.

Chow, S. -M., Lu, O., Cohn, J. F., & Messinger, D. S. (2017). Representing self-organization and non-stationarities in dyadic interaction processes using dynamic systems modeling techniques. In A. ovn Davier & P. Kyllonen (Eds.), *Innovative assessment of collaboration* (pp. 269–286). New York, NY: Springer.

Chun, C. A. (2016). The expression of posttraumatic stress symptoms in daily life: A review of experience sampling methodology and daily diary studies. *Journal of Psychopathology and Behavioral Assessment*, *38*(3), 406–420.

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, *57*, 505–528.

Cronin, M. A., & Vancouver, J. B. (2019). The only constant is change: Expanding theory by incorporating dynamic properties into one's models. In S. E. Humphrey & J. M. LeBreton (Eds.), *The handbook of multilevel theory, Measurement, and analysis* (pp. 89–114). Washington, DC: American Psychological Association.

Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp. 35–54). Dordrecht, The Netherlands: Springer Netherlands.

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62*, 583–619.

Depaoli, S., & Clifton, J. P. (2015). A Bayesian approach to multilevel structural equation modeling with continuous and dichotomous outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(3), 327–351.

Depaoli, S., & van de Schoot, R. (2017). Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological Methods, 22*, 240–261.

Epskamp, S., Deserno, M., & Bringmann, L. (2016). *mlVAR: multi-level vector autoregression*. R package version 0.3.3.

Epskamp, S., Waldorp, L. J., Mottus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research, 53*(4), 453–480.

Ferrer, E., & Nesselroade, J. R. (2003). Modeling affective processes in dyadic relations via dynamic factor analysis. *Emotion, 3*, 344–360.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subject research. *Proceedings of the National Academy of Sciences, 115*(27), E6106–E6115.

Fisher, C. D., & To, M. L. (2012). Using experience sampling methodology in organizational behavior. *Journal of Organizational Behavior*, *33*(7), 865–877.

Foster, K. T., & Beltz, A. M. (2018). Advancing statistical analysis of ambulatory assessment data in the study of addictive behavior: A primer on three person-oriented techniques. *Addictive Behaviors, 83*, 25–34.

Fraley, R. C., & Hudson, N. W. (2014). Review of intensive longitudinal methods: An introduction to diary and experience sampling research. *The Journal of Social Psychology, 154*(1), 89–91.

Gates, K. M., Fisher, Z. F., & Bollen, K. A. (2020). Latent variable GIMME using model implied instrumental variables (MIIVs). *Psychological Methods*, *25*(2),

227–242.

Gates, K. M., Lane, S. T., Varangis, E., Giovanello, K., & Guiskewicz, K. (2017). Unsupervised classification during time-series model building. *Multivariate Behavioral Research, 52*(2), 129–148.

Gates, K. M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *Neuroimage, 63*(1), 310–319.

Gottfredson, N. C., Panter, A. T., Daye, C. E., Allen, W. F., & Wightman, L. F. (2009). The effects of educational diversity in a national sample of law students: Fitting multilevel latent variable models in data with categorical indicators. *Multivariate Behavioral Research, 44*, 305–331.

Hamaker, E. L. (2017). *Dynamic structural equation modeling of intensive longitudinal data using Mplus Version 8: Parts 1 and 2.* Retrieved June 1, 2020, from https://www.statmodel.com/download/Aug17-18_JH_Slides.zip

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research, 53*(6), 820–841.

Hamaker, E. L., & Grasman, R. P. P. P. (2015). To center or not to center? Investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology, 5,* 1492.

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science, 26*, 10–15.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter.* Cambridge, UK: Cambridge University Press.

Haslbeck, J. M. B., & Waldorp, L. J. (2020). Mgm: Estimating time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software, 93*(8).

Hayes, S. C., Hofmann, S. G., Stanton, C. E., Carpenter, J. K., Sanford, B. T., Curtiss, J. E., & Ciarrochi, J. (2019). The role of the individual in the coming era of process-based therapy. *Behaviour Research and Therapy, 117*, 40–53.

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus.* (pp. 70–87). New York: Routledge.

Henry, T. R., Feczko, E., Cordova, M., Earl, E., Williams, S., Nigg, J. T., … Gates, K. M. (2019). Comparing directed functional connectivity between groups with confirmatory subgrouping GIMME. *Neuroimage, 188*, 642–653.

Hofmans, J., de Clercq, B., Kuppens, P., Verbeke, L., & Widiger, T. A. (2019). Testing the structure and process of personality using ambulatory assessment data: An overview of within-person and person-specific techniques. *Psychological Assessment, 31*(4), 432–443.

Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A comparison of ML, WLSMV, and Bayesian methods for multilevel structural equation models in small samples: A simulation study. *Multivariate Behavioral Research*, *51*(5), 661–680.

Howland, M., & Rafaeli, E. (2010). Bringing everyday mind reading into everyday life: Assessing empathic accuracy with daily diary data. *Journal of Personality*, *78*(5), 1437–1468.

Jahng, S. (2008). *Multilevel models for intensive longitudinal data with heterogeneous errors structure: Covariance transformation and variance function models*, M.A. Thesis. University of Missouri.

Jahng, S., & Wood, P. K. (2017). Multilevel models for intensive longitudinal data with heterogeneous autoregressive errors: The effect of misspecification and correction with Cholesky transformation. *Frontiers in Psychology, 8*, 262.

Johnson, J. E., Burlingame, G. M., Olsen, J. A., Davies, D. R., & Gleave, R. L. (2005). Group climate, cohesion, alliance, and empathy in group psychotherapy: Multilevel structural equation models. *Journal of Counseling Psychology, 52*, 310–321.

Jongerling, J., Laurenceau, J. P., & Hamaker, E. L. (2015). A multilevel AR (1) model: Allowing for inter-individual differences in trait-scores, inertia, and innovation variance. *Multivariate Behavioral Research, 50*, 334–349.

Kelava, A., & Brandt, H. (2019). A nonlinear dynamic latent class structural equation model. *Structural Equation Modeling-a Multidisciplinary Journal, 26*(4), 509–528.

Ke, Z., Zhang, Q., & Tong, X. (2019). Bayesian Meta-Analytic SEM: A one-stage approach to modeling between-studies heterogeneity in structural parameters. *Structural Equation Modeling-a Multidisciplinary Journal, 26*(3), 348–370.

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping, 28*(2), 85–93.

Kleiman, E. M., Glenn, C. R., & Liu, R. T. (2019). Real-time monitoring of suicide risk among adolescents: Potential barriers, possible solutions, and future directions. *Journal of Clinical Child and Adolescent Psychology, 48*(6), 934–946.

Lane, S. T., Gates, K. M., Fisher, Z., Arizmendi, C., & Molenaar, P. (2020). Gimme: Group iterative multiple model estimation. *Computer Software.* Retrieved April 14, 2020, from *https://CRAN.R-project.org/package=gimme*

Lane, S. T., Gates, K. M., Pike, H. K., Beltz, A. M., &

Wright, A. G. C. (2019). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods, 24*(1), 54.

Lanza, S. T., Vasilenko, S. A., Liu, X. Y., Li, R., & Piper, M. E. (2014). Advancing the understanding of craving during smoking cessation attempts: A demonstration of the time-varying effect model. *Nicotine & Tobacco Research, 16*, S127–S134.

Lau, Y., Tha, P. H., Wong, D. F. K., Wang, Y., Wang, Y., & Yobas, P. (2016). Different perceptions of stress, coping styles, and general well-being among pregnant Chinese women: A structural equation modeling approach. *Archives of Women's Mental Health, 19*(1), 71–78.

Lazarevic, L. B., Bjekic, J., Zivanovic, M., & Knezevic, G. (2020). Ambulatory assessment of language use: Evidence on the temporal stability of electronically activated recorder and stream of consciousness data. *Behavior Research Methods,* (1).

Li, R., Root, T. L., & Shiffman, S. (2006). A local linear estimation procedure of functional multilevel modeling. In T. Walls & J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 63–83). New York, NY: Oxford University Press.

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*(3), 203–229.

Malmberg, L. -E., & Martin, A. J. (2019). Processes of students' effort exertion, competence beliefs and motivation: Cyclic and dynamic effects of learning experiences within school days and school subjects. *Contemporary Educational Psychology*, *58*, 299–309.

McNeish, D. (2017). Challenging conventional wisdom for multivariate statistical models with small samples. *Review of Educational Research, 87*, 1117–1151.

McNeish, D. (2019). Two-level dynamic structural equation models with small samples. *Structural Equation Modeling-a Multidisciplinary Journal*, *26*(6), 948–966.

McNeish, D., & Hamaker, E. L. (2019, December). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*.

Miller, M. B., & van Horn, J. D. (2007). Individual variability in brain activations associated with episodic retrieval: A role for large-scale databases. *International Journal of Psychophysiology, 63*(2), 205–213.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika, 50*(2), 181–202.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2*(4), 201–218.

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18*, 112–117.

Molenaar, P. C. M., Rovine, M. J., & Corneal, S. E. (1999). Dynamic factor analysis of emotional dispositions of adolescent stepsons towards their stepfathers. In R. K. Silbereisen, & A. von Eye (Eds.). *Growing up in times of social change* (pp. 287–318). Berlin, Germany: de Gruyter.

Munsch, S., Meyer, A. H., Milenkovic, N., Schlup, B., Margraf, J., & Wilhelm, F. H. (2009). Ecological momentary assessment to evaluate cognitive-behavioral treatment for binge eating disorder. *International Journal of Eating Disorders, 42*(7), 648–657.

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (Eighth Edition). Los Angeles, CA: Muthén & Muthén.

Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 92–105). Washington, DC: American Psychological Association.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society, 49*(6), 1417–1426.

O'Brien, E., Coats, A., Owens, P., Petrie, J., Padfield, P. L., Littler, W. A., … Mee, F. (2000). Use and interpretation of ambulatory blood pressure monitoring: Recommendations of the British Hypertension Society. *British Medical Journal*, *320*(7242), 1128–1134.

Piccirillo, M. L., & Rodebaugh, T. L. (2019). Foundations of idiographic methods in psychology and applications for psychotherapy. *Clinical Psychology Review, 71*, 90–100.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233.

Rabe-Hesketh, S., Skrondal, A., & Zheng, X. (2007). Multilevel structural equation modeling. In S. -Y. Lee (Ed.), *Handbook of latent variable and related models* (pp. 209–227). Amsterdam, The Netherlands: Elsevier.

Racine, N., Plamondon, A., Hentges, R., Tough, S., & Madigan, S. (2019). Dynamic and bidirectional associations between maternal stress, anxiety, and social support: The critical role of partner and family support. *Journal of Affective Disorders, 252*, 19–24.

Reis, H. T., Gable, S. L., & Maniaci, M. R. (2014). Methods for studying everyday experience in its natural context. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd edition, pp. 373–403). New York: Cambridge University Press.

Schultzberg, M., & Muthén, B. (2018). Number of subjects and time points needed for multilevel time-series analysis:

A simulation study of dynamic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 25*(4), 495–515.

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods, 21*(2), 206–221.

Schuurman, N. K., Grasman, R. P. P. P., & Hamaker, E. L. (2016). A comparison of inverse-Wishart prior specifications for covariance matrices in multilevel autoregressive models. *Multivariate Behavioral Research, 51*(2-3), 185–206.

Schwartz, J. E., Stone, A. A., Shiffman, S., & Atienza, A. A. (2007). The analysis of real-time momentary data: A practical guide. In A. Stone, S. Shiffman, A. Arienza, & L. Nebeling (Eds.), *The science of real-time data capture: Self-reports in health research.* New York: Oxford University Press.

Sened, H., Lazarus, G., Gleason, M. E. J., Rafaeli, E., & Fleeson, W. (2018). The use of intensive longitudinal methods in explanatory personality research. *European Journal of Personality, 32*(3), 269–285.

Setodji, C. M., Martino, S. C., Dunbar, M. S., & Shadel, W. G. (2019, April 18). An exponential effect persistence model for intensive longitudinal data. *Psychological Methods. Advance Online Publication.24*(5), 622–636.

Shiffman, S. (2009). Ecological momentary assessment (EMA) in studies of substance use. *Psychological Assessment*, *21*(4), 486–497.

Shiyko, M. P., Li, Y., & Rindskopf, D. (2012). Poisson growth mixture modeling of intensive longitudinal data: An application to smoking cessation behavior. *Structural Equation Modeling: A Multidisciplinary Journal, 19*(1), 65–85.

Song, H., & Zhang, Z. (2014). Analyzing multiple multivariate time series data using multilevel dynamic factor models. *Multivariate Behavioral Research, 49*(1), 67–77.

Spencer, J. P., & Schöner, G. (2003). Bridging the representational gap in the dynamic systems approach to development. *Developmental Science*. *6*:392–412.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society Series B (Statistical Methodology), 64*, 583–616.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavorial medicine. *Annals of Behavioral Medicine, 16,* 199–202.

Stone, A. A., Shiffman, S. S., & DeVries, M. W. (1999). Ecological momentary assessment. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well being: The foundations of hedonic psychology* (pp. 26–39). New York, NY: Russell Sage Foundation.

Sturgeon, J. A., Zautra, A. J., & Arewasikporn, A. (2014). A multilevel structural equation modeling analysis of vulnerabilities and resilience resources influencing affective adaptation to chronic pain. *Pain*, *155*(2), 292–298.

Tennen, H., Affleck, G., Armeli, S., & Carney, M. A. (2000). A daily process approach to coping: Linking theory, research, and practice. *American Psychologist, 55*(6), 626–636.

Tong, X., & Zhang, Z. (2012). Diagnostics of robust growth curve modeling using student’s t distribution. *Multivariate Behavioral Research, 47*(4), 493–518.

Trull, T. J., & Ebner-Priemer, U. W. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology, 9,* 151–176.

Trull, T. J., & Ebner-Priemer, U. W. (2014). The role of ambulatory assessment in psychological science. *Current Directions in Psychological Science, 23*(6), 466–470.

Trull, T. J., Solhan, M. B., Tragesser, S. L., Jahng, S., Wood, P. K., Piasecki, T. M., & Watson, D. (2008). Affective instability: Measuring a core feature of borderline personality disorder with ecological momentary assessment. *Journal of Abnormal Psychology, 117*(3), 647–661.

Turner, A. E., & Hayes, S. C. (1996). A comparison of response covariation viewed idiothetically and nomothetically. *Psicologia Conductual, 4*, 231–250.

Vallacher, R. R., Read, S. J., & Nowak, A. (2002). The dynamical perspective in personality and social psychology. *Personality and Social Psychology Review*, *6*, 264–273.

van de Schoot, R., Broere, J. J., Perryck, K. H., Zondervan-Zwijnenburg, M., & van Loey, N. E. (2015). Analyzing small data sets using Bayesian estimation: The case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *European Journal of Psychotraumatology, 6*, 25216.

van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development, 85,* 842–860.

van Roekel, E., Keijsers, L., & Chung, J. M. (2019). A review of current ambulatory assessment studies in adolescent samples and practical recommendations. *Journal of Research on Adolescence*, *29*(3), 560–577.

Voelkle, M. C., Gische, C., Driver, C. C., & Lindenberger, U. (2018). The role of time in the quest for understanding psychological mechanisms. *Multivariate Behavioral Research, 53*(6), 782–805.

Walls, T. A., & Schafer, J. L. (2006). *Models for intensive longitudinal data.* New York, NY: Oxford University

Press.

Walls, T. A., Hoppner, B., & Goodwin, M. (2007). Statistical issues in intensive longitudinal data analysis. In A. A. Stone, S. Shiffman, A. A. Atienza, & L. Nebeling (Eds.), *The science of real-time data capture: Self-reports in health research* (pp. 338–360). New York, NY: Oxford University Press.

Weinstock, J., Farney, M. R., Elrod, N. M., Henderson, C. E., & Weiss, E. P. (2017). Exercise as an adjunctive treatment for substance use disorders: Rationale and intervention description. *Journal of Substance Abuse Treatment, 72*, 40–47.

West, S. G. (2018). Opportunities and issues in modeling intensive longitudinal data: Learning from the COGITO project. *Multivariate Behavioral Research*, *53*(6), 777–781.

Windt, J., Ardern, C. L., Gabbett, T. J., Khan, K. M., Cook, C. E., Sporer, B. C., & Zumbo, B. D. (2018). Getting the most out of intensive longitudinal data: A methodological review of workload-injury studies. *BMJ Open, 8*(10), e022626.

Woods, W. C., Arizmendi, C., Gates, K. M., Stepp, S. D., Pilkonis, P. A., & Wright, A. G. C. (2020). Personalized models of psychopathology as contextualized dynamic processes: An example from individuals with borderline personality disorder. *Journal of Consulting and Clinical Psychology, 88*(3), 240–254.

Wright, A. G. C., Hallquist, M. N., Stepp, S. D., Scott, L. N., Beeney, J. E., Lazarus, S. A., & Pilkonis, P. A. (2016). Modeling heterogeneity in momentary interpersonal and affective dynamic processes in borderline personality disorder. *Assessment, 23*, 484–495.

Wright, A. G. C., Hopwood, C. J., & Simms, L. J. (2015). Daily interpersonal and affective dynamics in personality disorder. *Journal of Personality Disorders, 29*, 503–525.

Wright, A. G. C., & Woods, W. C. (2020). Personalized models of psychopathology. *Annual Review of Clinical Psychology, 16*, 49–74.

Zhang, J., & Zheng, Y. (2019). Neuroticism and extraversion are differentially related to between- and within-person variation of daily negative emotion and physical symptoms. *Personality and Individual Differences, 141*, 138–142.

Zhang, Z. Y. (2013). Bayesian growth curve models with the generalized error distribution. *Journal of Applied Statistics, 40*(8), 1779–1795.

Zhang, Z. Y., Lai, K., Lu, Z., & Tong, X. (2013). Bayesian inference and application of robust growth curve models using student's t distribution. *Structural Equation Modeling: A Multidisciplinary Journal, 20*(1), 47–78.

Zhang, Z. Y., & Nesselroade, J. R. (2007). Bayesian estimation of categorical dynamic factor models. *Multivariate Behavioral Research, 42*(4), 729–756.

Zhou, L., Wang, M., & Zhang, Z. (2020). Intensive longitudinal data analyses with dynamic structural equation modeling. *Organizational Research Methods*.

Zondervan-Zwijnenburg, M., Peeters, M., Depaoli, S., & van de Schoot, R. (2017). Where do priors come from? Applying guidelines to construct informative priors in small sample research. *Research in Human Development, 14*, 305–320.

Intensive longitudinal data analysis: Models and application

ZHENG Shufang, ZHANG Lijin, QIAO Xinyu, PAN Junhao

(*Department of Psychology, Sun Yat-sen University, Guangzhou* 510006, *China*)

Abstract: In the fields of psychology, education, and clinical science, researchers have devoted increased attention to the dynamic changes and personalized modeling of individuals' behaviors, minds, and treatment effects over time. Intensive longitudinal data is a set of measures collected at multiple time points with higher frequency over shorter periods. Thus, it can be used in the analysis of the dynamics and mechanisms of within-person processes. In recent years, intensive longitudinal design has become one of the most prominent and promising approaches in psychological research. However, many of these researches still rely on traditional data analysis methods. Many models have been proposed to analyze intensive longitudinal data, including top-down approaches (e.g., dynamic structural equation model, DSEM) and bottom-up approaches (e.g., group iterative multiple model estimation, GIMME). Both of the methods can conveniently model autoregressive and cross-lagged effects in intensive longitudinal data.

Key words: intensive longitudinal data, time-series, DSEM, GIMME

附录 A：MLM 与 RDSEM 实证演示 *Mplus* 代码

1. MLM 代码：

```
DATA:
FILE = [DATA LOCATION PATH]\ two-level data with Trend.csv;
  VARIABLE:
    NAMES ARE URGE DEP JS HS PERSON TIME;
    USEVARIABLES URGE DEP JS HS TIME;
    cluster=PERSON;
    BETWEEN = JS HS;
    WITHIN = TIME DEP;! Time is only used in the Within-Level model;

  DEFINE:
    CENTER DEP (groupmean);

  ANALYSIS:
    TYPE = TWOLEVEL RANDOM;!Two-Level Model with paths that are latent variables;

  MODEL:
    %WITHIN%
    b1i | URGE ON DEP;
    ! urge to smoke is regressed on Lag-1 urge to smoke
    trend | URGE ON TIME;! urge is regressed on time to capture the trend over time, the !slope is latent;

    %BETWEEN%
    URGE b1i;
    URGE with b1i TREND;b1i with TREND;

    [URGE]; !mean intercept of urge to smoke;
    [b1i]; !mean of the regressive slope;
    [TREND];

    URGE; !urge to smoke intercept variance;
    b1i; !regressive slope variance;
    TREND;

    URGE on JS HS;
    !The person-specific intercept is predicted by Job Stress and Home Stress;
    b1i on JS HS;
    !The regressive slope is predicted by Job Stress and Home Stress;
    TREND on JS HS;

  OUTPUT: TECH1 TECH8;
```

2. RDSEM 代码：

```
DATA:
FILE = [DATA LOCATION PATH]\ two-level data with Trend.csv;
  VARIABLE:
    NAMES ARE URGE DEP JS HS PERSON TIME;
    USEVARIABLES URGE DEP JS HS TIME;
    Lagged = URGE (1) DEP (1);!Created a Lag-1 variable for urge to smoke and depression;
    CLUSTER=PERSON;
    BETWEEN = JS HS;! Job Stress and Home Stress are only used in the Between-Level model;
    WITHIN = TIME;! Time is only used in the Within-Level model;

  ANALYSIS:
    TYPE = TWOLEVEL RANDOM;!Two-Level Model with paths that are latent variables;
    ESTIMATOR = BAYES;!Lagged variables can only be estimated with Bayes in Mplus;
    BITERATIONS = (1000);!Run at least 1000 iteration of the MCMC algorithm;
    BSEED=1028;

  MODEL:
    %WITHIN%
    PHI1 | URGE^ ON URGE^1;
    ! urge to smoke is regressed on Lag-1 urge to smoke
    PHI2 | DEP^ ON DEP^1;
    ! depression is regressed on Lag-1 depression
    PHI3 | DEP^ ON URGE^1;
    ! Depression is regressed on Lag-1 urge to
```

```
smoke
        PHI4 | URGE^ ON DEP^1;
        ! urge to smoke is regress on Lag-1 depression
        TREND1 | URGE ON TIME;! urge is regressed
on time to capture the trend over time
        TREND2 | DEP ON TIME;! dep is regressed on
time to capture the trend over time
        LogV1 | Urge;
        ! the Within-Level residual variance of urge to
smoke
        LogV2 | DEP;
        ! the Within-Level residual variance of
depression

        %BETWEEN%
        [URGE]; !mean intercept of urge to smoke
        [DEP]; !mean intercept of depression
        [PHI1]; !mean of urge autoregressive slope
        [PHI2]; !mean of depression autoregressive
slope
        [PHI3]; !mean of dep on urge&1 slope
        [PHI4]; !mean of urge on dep&1 slope
        [TREND1];
        [TREND2];
        [LogV1];
        !log of the mean of the urge residual variance
        [LogV2];
        !log of the mean of the urge residual variance

        URGE; !urge to smoke intercept variance
        DEP; !depression intercept variance
        PHI1; !urge to smoke autoregressive slope
variance
        PHI2; !depression autoregressive slope variance
        PHI3; !dep on urge&1 slope variance
        PHI4; !urge on dep&1 slope
        TREND1;
        TREND2;
        LogV1;
        !between-person variance of urge to smoke
residual variance
        LogV2; !between-person variance of depression
residual variance
        URGE on JS HS;
        DEP on JS HS;
        !The person-specific intercepts are predicted by
Job Stress and Home Stress;
        PHI1 on JS HS;
        PHI2 on JS HS;
        PHI3 on JS HS;
        PHI4 on JS HS;
        !The person-specific slopes are predicted by Job
Stress and Home Stress;
        LOGV1 on JS HS;
        LOGV2 on JS HS;
        !The person-specific residual variances are
predicted by Job Stress and Home Stress;

        URGE DEP PHI1-PHI4 LOGV1 LOGV2 WITH
URGE DEP PHI1-PHI4 LOGV1 LOGV2;

    OUTPUT: TECH1 TECH8 STDYX;
```

附录 B：GIMME 实证演示 R 代码

```
library (DiagrammeR)
library (gimme)
library (mvtnorm)
library (dplyr)
library (xlsx)

#1、数据预处理
wd<-"~" # 设置文件夹路径
setwd (wd)
dat<- read.csv ("Two-Level Data with Trend.csv",
header=F) # 读取数据文件
colnames (dat)<-c ("URGE", "DEP", "JS", "HS",
"PERSON", "TIME") # 指定各变量名，其中 URGE 指
吸烟欲望，DEP 指抑郁水平
datawd<-paste0(wd, "/data")
dir.create(datawd)
setwd (datawd)
for (i in 1:length (unique (dat$PERSON))){
# 选取个体 i 的数据进行预处理
person_dat<-dat[which (dat$PERSON==i), c ("URGE",
```

```
"DEP", "TIME")]

#分别去除变量 URGE 和 DEP 随时间变化的整体趋势(detrend)
temp = lm (person_dat$URGE~person_dat$TIME)
person_dat$urge<-residuals (temp)
temp2 = lm (person_dat$DEP~person_dat$TIME)
person_dat$dep<-residuals (temp2)

#将个体 i 的数据单独写入 excel 文档"data i.csv"
write.csv (person_dat[, c ("urge", "dep")],
  paste ("data", i, ".csv"), row.names = FALSE)
}

#2、利用 gimmeSEM 函数输出群体和个体模型结果
fit<-gimmeSEM (
  data = datawd,  # 存放数据的文件夹路径
  out = paste0(wd, "/output"),  # 存放输出结果的文件夹路径
  sep = ", ",  # 声明数据分隔符，这里 csv 格式的数据分隔符为逗号
  header = T,  # 声明数据中是否包括变量名
  ar = TRUE,  # 指定模型中是否包括自回归，默认为 TRUE
  plot = TRUE,  # 指定结果是否输出模型图，默认为 TRUE
  groupcutoff =.75  # 路径可以被纳入群体模型的临界值(显著改善个体模型的最小比例，默认为 75%)
    )
fit2<-aggSEM(
  data = datawd,  # 存放数据的文件夹路径
  out = paste0(wd,'/output'), # 存放输出结果的文件夹路径
  sep = ",",
  header = T)
```